ESSENTIALISM: PARADISE LOST

GEORGE DJUKIC

DEPARTMENT OF PHILOSOPHY

UNIVERSITY OF ADELAIDE

PH.D. THESIS

COMPLETED FOR SUBMISSION AUGUST 1997

FOR MY FATHER,

WHO TAUGHT ME TO LOVE WORDS,

AND

FOR MICHAEL BRADLEY,

WHO TAUGHT ME TO TAME THAT LOVE.

1) **THIS THESIS CONTAINS NO MATERIAL WHICH HAS BEEN ACCEPTED FOR THE AWARD OF ANY OTHER DEGREE OR DIPLOMA IN ANY UNIVERSITY.**

2) **TO THE BEST OF MY KNOWLEDGE AND BELIEF, THIS THESIS CONTAINS NO MATERIAL PREVIOUSLY PUBLISHED OR WRITTEN BY ANOTHER PERSON, EXCEPT WHERE DUE REFERENCE HAS BEEN MADE IN THE TEXT.**

3) **I GIVE CONSENT TO THIS COPY'S BEING AVAILABLE FOR LOAN AND PHOTOCOPYING ONCE IT HAS BEEN DEPOSITED IN THE UNIVERSITY LIBRARY.**

SIGNED

DATE

ABSTRACT OF THESIS

This thesis argues for the incoherency of the putative distinction between those properties which are possessed essentially - in a **metaphysical** sense - by an object and those which are possessed only accidentally. This thesis is thus a critique of **metaphysical** essentialism.

In chapter one I distinguish various kinds of essentialism, possibility and necessity. I explore both metaphysical and semantical issues associated with the essential/accidental distinction. I examine Quine's related attack on quantified modal logic. The burden of Quine's objections is carried by his antipathy to metaphysical essentialism, arguments against which he leaves undeveloped.

In chapter two I remedy this defect and develop a line of attack on the essential/accidental distinction by adapting and sytematising an argument of Chisholm's in his seminal paper of 1967. I consider various lines of response by the essentialist and find them wanting. I concede that the weakest point in my attack on essentialism is an appeal to a principle about the transworld identity conditions of individuals.

Given this concession, my discussion changes tack in chapters three and four. I grant irenically the coherence of the essential/accidental distinction. In chapter three I argue that even given this concession there is no cogent case to be made for the claim commonly advanced by essentialists that an object's origin is essential to that object's identity. What I take to be the stongest argument for origin essentialism is in effect self-refuting because it appeals to the very principle of transworld identity which grounds my rejection of essentialism in chapter two.

In chapter four I argue that there is no compelling ground to believe that an object essentially satisfies - in a **metaphysical** as opposed to a conceptual sense of 'essentially' - the sortals which are true of it.

CONTENTS

CHAPTER 1 ESSENTIALISM, QUINE AND QML

1.1	INTRODUCTION	3
1.2	ESSENTIALISM AND ANTI-ESSENTIALISM	6
1.2.1	LOGICAL, CAUSAL AND METAPHYSICAL POSSIBILITY AND IMPOSSIBILITY	6
1.2.2	WOULD HAVE AND COULD HAVE	12
1.2.3	METAPHYSICAL ESSENTIALISM VERSUS CONCEPTUAL ESSENTIALISM	15
1.2.4	FORMAL AND ORDINARY LANGUAGE CHARACTERISATIONS OF ESSENTIALISM	19
1.3	INFORMAL MODAL SEMANTICS	21
1.3.1	NECESSARILY TRUE THAT P	21
1.3.2	NECESSARILY IS F	23
1.3.2.1	(∃X)(NECF)X AND (∃X)NEC(FX)	28
1.3.2.2	(X)(NECF)X AND (X)NEC(FX)	30
1.3.3	EQUIVALENCE OF '(NECF)A' AND 'NEC(FA)'	30
1.3.3.1	EMPTY NAMES 'A'	30
1.3.3.2	RUSSELLIAN INTERPRETATION OF DEFINITE DESCRIPTIONS	31
1.3.3.3	A PARTICULAR NON-RUSSELLIAN INTERPRETATION OF DEFINITE DESCRIPTIONS	32
1.3.3.4	INTERPRETING 'NECG((IX)FX)', WHERE '(IX)FX' IS CONSTRUED AS A SINGULAR TERM	35
1.3.4	CONTINGENT TRUTHS MAY BE UNDERSTOOD AS NECESSARY, AND CONVERSELY	36

1.3.5 SUMMARY OF THE MODAL SEMANTICS 38
1.4 DO QUINE'S FORMAL AND INFORMAL CHARACTERISATIONS OF ESSENTIALISM COINCIDE? 39
1.5 QUINE ON THE ILLEGITIMACY OF QML 41
1.5.1 QUANTIFICATION INTO MODAL CONTEXTS 42
1.5.2 THE ARGUMENT FOR UNINTELLIGIBILITY FROM THE REFERENTIAL OPACITY OF QUANTIFICATIONAL CONTEXTS 44
1.5.3 THE ARGUMENT FOR UNINTELLIGIBILITY FROM THE INDETERMINACY OF TRUTH-VALUE OF QUANTIFIED MODAL CONTEXTS 50
1.5.3.1 RESPONSE TO THE ARGUMENT FROM THE INDETERMINACY OF TRUTH-VALUES 53
1.5.4 IS QML COMMITTED TO ESSENTIALISM? 60
1.5.5 QUINE'S OBJECTIONS TO ESSENTIALISM 62
1.5.5.1 ESSENTIALISM CONFLICTS WITH THE EXPLANATION OF NECESSITY THROUGH ANALYTICITY 62
1.5.5.2 ESSENTIALISM LEADS TO INCONSISTENCY 68
APPENDIX A 71

CHAPTER 2 THE PROBLEM FOR ESSENTIALISM OF EXISTENCE IN MANY WORLDS

2.1 THE COHERENCY OF THE NOTION OF EXISTENCE IN MANY WORLDS 77
2.1.1 COMMENTS ON CHISHOLM'S FORMULATION 82

2.2	CHISHOLM'S ARGUMENT FROM THE POINT OF VIEW OF MODAL LOGICS	89
2.2.1	THE NECESSITY OF DISTINCTNESS	90
2.2.2	THE NECESSITY OF IDENTITY	93
2.2.2.1	AN ESSENTIALIST REPLY	94
2.2.2.2	ASSESSING THE ESSENTIALIST'S REPLY	96
2.2.3	THE NECESSITY OF DISTINCTNESS	100
2.3	AN ESSENTIALIST REPLY	103
2.3.1	REJECT THE IDENTITY OF INDISCERNIBLES FOR COUNTERFACTUAL POSSIBILITY	103
2.3.2	REJECT THE POSSIBILITY PRINCIPLE	110
2.3.2.1	HAECCEITIES AS UNIQUELY ESSENTIAL	114
2.3.2.2	THE ESSENTIALITY OF UNIQUE WORLD-INDEXED PROPERTIES	120
2.3.3	REJECT THE PRINCIPLE OF THE TRANSITIVITY OF PREDICATIVE POSSIBILITY	123
2.3.4	ARE KRIPKEAN STIPULATIONS A SOLUTION?	126
2.4	LOOKING AHEAD TO CHAPTER 3	128

CHAPTER 3 THE ESSENTIALITY OF ORIGIN

3.1	PRELIMINARIES	131
3.2	A COULD BE F; A COULD HAVE BEEN F; A COULD HAVE BECOME F	133
3.3	J.L.MACKIE ON ORIGIN	139
3.4	THE ARGUMENT FROM THE BETTER CANDIDATE	143
3.5	REFERENCE TO FUTURE INDIVIDUALS	144
3.6	MCGINN ON THE ORIGIN OF ORGANISMS	147

3.6.1 IDENTITY OF ZYGOTE WITH ADULT 147
3.6.1.1 ZYGOTE AS DISTINCT FROM ITS CONSTITUTIVE CELL 152
3.6.1.2 THE POSSIBILITY OF SLIGHT DIFFERENCE IS INCOMPATIBLE WITH THE IMPOSSIBILITY OF COMPLETE DIFFERENCE 154
3.6.1.2.1 THE CASE OF MATERIAL COMPOSITION 154
3.6.1.2.2 THE CASE OF GENETIC STRUCTURE 157
3.6.2 GAMETES AND THEIR PARENTS 158
3.6.3 THE ESSENTIALITY OF GAMETES 159
3.7 KRIPKEAN ARGUMENTS FOR THE ESSENTIALITY OF ORIGINAL MATERIAL COMPOSITION 160
3.7.1 KRIPKE'S ARGUMENT 160
3.7.2 SALMON'S ARGUMENT 1 161
3.7.3 SALMON'S ARGUMENT 2 162
3.7.3.1 A PROBLEM FOR V** 166
3.7.3.2 A PROBLEM FOR NOONAN'S PROPOSAL 169
3.7.3.3 EVALUATING NOONAN'S PROPOSAL 172
3.8 AN ARGUMENT AGAINST THE ESSENTIALITY OF ORIGIN 173
3.9 FORBES ON THE ESSENTIALITY OF ORIGIN 176

CHAPTER 4 THE ESSENTIALITY OF SORTAL SATISFACTION

4.1 INTRODUCTION 188
4.2 UNITY AT A TIME AND OVER TIME 189
4.3 SORTALS 191
4.4 SORTALS, NATURES AND NATURAL KINDS 199
4.5 SORTALS AND CAUSALITY 204
4.6 PUTNAM ON THE ESSENTIALITY OF SORTALS 210

4.7 BRODY ON THE ESSENTIALITY OF SORTALS 219

4.8 WIGGINS ON THE ESSENTIALITY OF SORTALS 222

4.9 THE OMNITEMPORALITY THESIS 228

GLOSSARY 238

BIBLIOGRAPHY 243

CONTENTS

CHAPTER 1 ESSENTIALISM, QUINE AND QML

1.1	INTRODUCTION	3
1.2	ESSENTIALISM AND ANTI-ESSENTIALISM	6
1.2.1	LOGICAL, CAUSAL AND METAPHYSICAL POSSIBILITY AND IMPOSSIBILITY	6
1.2.2	WOULD HAVE AND COULD HAVE	12
1.2.3	METAPHYSICAL ESSENTIALISM VERSUS CONCEPTUAL ESSENTIALISM	15
1.2.4	FORMAL AND ORDINARY LANGUAGE CHARACTERISATIONS OF ESSENTIALISM	19
1.3	INFORMAL MODAL SEMANTICS	21
1.3.1	NECESSARILY TRUE THAT P	21
1.3.2	NECESSARILY IS F	23
1.3.2.1	(∃X)(NECF)X AND (∃X)NEC(FX)	28
1.3.2.2	(X)(NECF)X AND (X)NEC(FX)	30
1.3.3	EQUIVALENCE OF '(NECF)A' AND 'NEC(FA)'	30
1.3.3.1	EMPTY NAMES 'A'	30
1.3.3.2	RUSSELLIAN INTERPRETATION OF DEFINITE DESCRIPTIONS	31
1.3.3.3	A PARTICULAR NON-RUSSELLIAN INTERPRETATION OF DEFINITE DESCRIPTIONS	32
1.3.3.4	INTERPRETING 'NECG((IX)FX)', WHERE '(IX)FX' IS CONSTRUED AS A SINGULAR TERM	35
1.3.4	CONTINGENT TRUTHS MAY BE UNDERSTOOD AS NECESSARY, AND CONVERSELY	36

1.3.5	SUMMARY OF THE MODAL SEMANTICS	38
1.4	DO QUINE'S FORMAL AND INFORMAL CHARACTERISATIONS OF ESSENTIALISM COINCIDE?	39
1.5	QUINE ON THE ILLEGITIMACY OF QML	41
1.5.1	QUANTIFICATION INTO MODAL CONTEXTS	42
1.5.2	THE ARGUMENT FOR UNINTELLIGIBILITY FROM THE REFERENTIAL OPACITY OF QUANTIFICATIONAL CONTEXTS	44
1.5.3	THE ARGUMENT FOR UNINTELLIGIBILITY FROM THE INDETERMINACY OF TRUTH-VALUE OF QUANTIFIED MODAL CONTEXTS	50
1.5.3.1	RESPONSE TO THE ARGUMENT FROM THE INDETERMINACY OF TRUTH-VALUES	53
1.5.4	IS QML COMMITTED TO ESSENTIALISM?	60
1.5.5	QUINE'S OBJECTIONS TO ESSENTIALISM	62
1.5.5.1	ESSENTIALISM CONFLICTS WITH THE EXPLANATION OF NECESSITY THROUGH ANALYTICITY	62
1.5.5.2	ESSENTIALISM LEADS TO INCONSISTENCY	68
	APPENDIX A	71

CHAPTER 1: ESSENTIALISM, QUINE AND QML

1.1 INTRODUCTION

Hilbert once said of Cantor, 'From the Paradise created by Cantor no one can ever drive us.' I aim to show that the same does not hold of possible worlds. Possible worlds have been called 'a philosopher's paradise'. In the past two decades and more what I shall call 'metaphysical essentialism' - to be contrasted with what I shall call 'conceptual essentialism' - has thrived in symbiosis with possible worlds in philosophers' paradise. In this thesis I aim to drive these parasites out of paradise, lest they turn it into a metaphysical jungle.

My point of departure is the notion of an essential property. On one view, the Aristotelian, an essential property is part of an object's essence. The essence of an object makes the object what it is (whatever that is supposed to mean exactly). In modern debates, however, the notion of an essential property is primarily a conception of the mode in which a property is possessed by its bearer. Of course philosophers have wanted to draw connections between both concepts. Nevertheless I want to avoid possible ambiguity and restrict my understanding of the concept of an essential property to the second conception. On this view, a property of an object is an essential property of that object just where it is a property which is necessarily possessed by its bearer. I hereby name any theory (doctrine, account, view, thesis, proponent, etc.) which (who) admits or presupposes the coherency of a distinction between, on the one hand, essential (necessary) properties and, on the other hand, non-essential (non-necessary or accidental or contingent) properties, 'an essentialist theory (doctrine, account, view, thesis, proponent, etc.)'. As Cartwright explains, an object's essential properties are those it has necessarily, those it could not have lacked, whereas its accidental properties are those it has only contingently, those it might not have had. (Cartwright, p.149.) In other words, an object a has a property essentially if and

only if any object lacking that property **would** not be or have been a. (So we could say that the property in question is essential to a's identity.) This is of course a different condition from the condition that any object lacking a certain property **is** not a. On the second condition, for example, the property of being made of steel - as opposed to a steel/titanium alloy - is held to be essential to the Eiffel tower simply because any actual tower lacking that property is, as a matter of empirical fact, not the Eiffel tower. What the **first** condition is intended by essentialists to rule in as an essential property of the Eiffel tower is, for example, its being a **tower** - whether steel or steel/titanium alloy - as opposed to its being a crane - whether steel or steel/titanium alloy. Thus any future change, whether actual or non-actual, that results in the meltdown of the Eiffel tower and its subsequent recasting as a crane, will not preserve, so the essentialist will say, the existence of the Eiffel tower. Similarly, any past change (in this case non-actual), having similar results to those of the future change of the previous sentence, would not have preserved the tower's existence. Moreover, any complex of past changes (again, as it turns out, non-actual) that prevented the construction of any towers whatsoever in the earth's history would have disqualified the identification of the Eiffel tower with **every** object that would have come into existence. These sorts of things are intended by metaphysical essentialism, i.e the view that endorses the **coherency** of the notion that a property F is **metaphysically essential** to an object a.

Contrast metaphysical essentialism with what I shall call 'conceptual essentialism'. This is the view that endorses the coherency of the notion that an object has **conceptually necessary** properties. According to conceptual essentialism, a property F is conceptually essential to an object a just where a's being F is **logically entailed** by a's being G. For example, John's being taller than Mary is conceptually essential or necessary to John given the fact, say, that Mary is shorter than John. Similarly, it is stipulated by Eastern Orthodox church canon that clerics of the

Eastern Orthodox be unmarried if they are to qualify for the patriarchate; one could say therefore that being unmarried is conceptually essential to being a patriarch of the Orthodox church. It follows that there will be properties which will be held by some to be conceptually essential but not metaphysically essential, or both conceptually and metaphysically essential, or neither conceptually nor metaphysically essential. Thus it might be urged that being unmarried is conceptually but not metaphysically essential to the patriarch. It is my contention that metaphysical essentialism is philosophically interesting but incoherent, while conceptual essentialism is coherent, indeeed true, but philosophically uninteresting.

For the purposes of this thesis I shall understand 'object'in its paradigmatic sense, i.e. that of 'concrete object', objects in space-time. (Thus I do not consider persons *qua* persons but *qua* spatio-temporal objects, if there is such a distinction to be made.) This of course is not to rule in advance that abstract objects cannot have essential properties; rather, it is to point out that my primary focus in this thesis will be on objects in the paradigmatic sense, and furthermore on **actual** as opposed to non-actual objects. Given my doubts about the coherency of the notion of metaphysical possibility, understandably I should want to avoid the introduction of non-actual individuals, although they will be called on in my discussion of Chisholm's argument in chapter 2.

Now among other things Quine is famous for his unremitting campaign against quantified modal logic (henceforth QML), a campaign which attains historical importance for its prompting of a resurgence of philosophical interest in the latter half of this century in the question of essential versus accidental properties. Quine's principal contention is that systems of QML are committed to what he calls 'Aristotelian essentialism', a doctrine he considers suspect, and that accordingly systems of QML inherit this suspectness and are best proscribed. Quine's arguments raise a welter of entangled issues. It is my aim in this chapter to isolate

some of these issues and to disentangle the more prominent lines of argument bearing on the issue of essentialism.

1.2 ESSENTIALISM AND ANTI-ESSENTIALISM

1.2.1 LOGICAL,CAUSAL AND METAPHYSICAL POSSIBILITY AND IMPOSSIBILITY

The principal sources for Quine's arguments are his four papers *Reference and Modality* (RM), *Three Grades of Modal Involvement* (3GMI), *Worlds Away* (WA), and *Reply to Professor Marcus* (RPM), and excerpts from his *Word and Object* (WO). But before I can examine Quine's arguments it will be useful to explicate certain logical and metaphysical concepts crucial to an understanding of the various issues involved. Quine himself has little to say about what exactly he means by 'Aristotelian essentialism'. He says at one place that it is the tradition whose origins are standardly attributed to Aristotle, which recognises a distinction between the necessary (essential) and contingent (accidental) attributes of an object, between its essence and accidents. (Quine WO, p.199.) Elaborating briefly, he explains that according to this tradition, an 'object of itself and by whatever name or none [i.e. independently of the language in which it is referred to, if at all (cf. Quine 3GMI, pp. 175-76)] must be seen as having some of its traits necessarily and others contingently, despite the fact that the latter traits follow just as analytically from some ways of specifying the object as the former traits do from other ways of specifying it.' (Quine RM, p.148.) Quine also provides a more formal characterisation of Aristotelian essentialism:

> More formally, what Aristotelian essentialism says is that you can have open sentences - which I shall here represent as 'Fx' and 'Gx' - such that ... $(\exists x)(\text{nec}Fx\&Gx\&\sim\text{nec}Gx)$. (Quine 3GMI, p.176.)

Unfortunately, this deceptively straightforward formal characterisation houses a nest

of logical issues whose tracing out will require a considerable digression.

Aristotle had of course a theory of essential as opposed to accidental properties, but Quine's remarks as they stand are insufficient to distinguish Aristotle's theory of essential properties from some non-Aristotelian theories. (Quine of course leaves it open whether his characterisation of Aristotelian essentialism in fact conforms to Aristotle's own theory of essence. For an instructive discussion of what Aristotle may have had in mind, which is critical of various assumptions common to some modern accounts of Aristotle's theory, see Matthews 1990. Cohen's fascinating and brilliant paper of 1977-78 in the *Review of Metaphysics* is also valuable reading.)

Some terminology (see the glossary at the end of this thesis for introduced terminology and acronyms): I shall understand '**ESSENTIALISM**' as the thesis that the distinction between an object's essential and accidental properties is a coherent one. A proponent of essentialism will be called 'an **ESSENTIALIST**'. '**FULL-ESSENTIALISM**' I shall understand as the thesis that some of an object's properties are in reality essential to it.. Clearly it does not follow from my definition of full-essentialism that an object need have any accidental properties at all. In fact we can discriminate between various forms of full-essentialism. What I shall call '**CATHOLIC ESSENTIALISM**' holds that all objects must have the same essential properties; what I shall call '**INDIVIDUALIST ESSENTIALISM**' holds that not every one of a particular object's essential properties need be essential to another object, so that individualist essentialism is inconsistent with catholic essentialism. It follows from my definitions that one can consistently defend an essentialist theory - i.e. one can be an essentialist - without wanting to defend full-essentialism. Thus one might recognise as legitimate the distinction between essential and accidental properties while believing that in reality no object has any of

its properties necessarily. On the other hand a proponent of full-essentialism, a **FULL-ESSENTIALIST, is** committed *inter alia* to the coherency of the essential/accidental distinction and so must support an essentialist theory of one kind or other. An opponent of full-essentialism, the **ANTI-FULL-ESSENTIALIST**, may, as we have just observed, still endorse some form of essentialist theory, but need not. Quine for example is an anti-full-essentialist not so much because he thinks that in reality no property is essential to anything, but because he finds incoherent the very distinction, fundamental to full-essentialism, between essential and accidental properties, i.e. he is also an **ANTI-ESSENTIALIST**. (In fact, he would probably grant that he is at most a conceptual essentialist, but with reservations of course, given his doubts about analyticity.) I shall examine Quine's concerns in due course.

But *prima facie* at least it is not clear that this distinction is incoherent. After all, unless one is a strict determinist, one is typically willing - pre-analytically anyway - to grant that things could have been otherwise. The question then arises: to what **extent** could things have been otherwise? Are there respects in which things could not have been otherwise? For their part, a strict determinist, i.e. one who holds both that all events have causes and that causes necessitate their effects, would reply that things cannot be or could not have been otherwise. Nevertheless they might allow that things could logically have been otherwise, so that for example it is logically possible - simply because it is not self-contradictory - to assert that Napoleon won the Battle of Waterloo, even though he could not have won it in some causal sense of 'could'. If asked to explain this notion of causal impossibility the strict determinist would probably appeal to the notion of causal necessitation: i.e. a cause necessitates its effect, so that given the events leading up to the Battle of Waterloo Napoleon had to lose. As for this notion of having to lose I do not know how to

explicate it without circularity; instead I assume the reader is acquainted with the pre-analytic intuitions which inform it.

For their part, an opponent of strict determinism, who nonetheless shared the necessitarian view of the relation between cause and effect, would typically maintain that things could have been otherwise, in both a causal and logical sense. Take for instance Napoleon. Napoleon's defeat according to them was not causally necessary. For at some time before the battle the conjunction of conditions sufficient for Naponeon's defeat had not yet obtained; moreover, some of these as-yet-to-obtain conditions were themselves not predetermined, so that while they eventually obtained they did not have to obtain, because they were eventually determined by events which did not have to occur and therefore by events which would have had no causes - hence the indeterminism. Again, I do not know how to explicate this notion of not having to obtain or occur, without appealing to the reader's acquaintance with the pre-analytic intuitions which underlie this notion. In terms of the metaphor of a replay, the opponent of strict determinism might say that in a replay of events leading up to Waterloo, Napoleon wins. But this is a metaphor.

We recall that the strict determinist would probably want to allow that Napoleon's winning Waterloo is logically possible although causally impossible, while the opponent of strict determinism would probably want to urge that this state of affairs is both logically and causally possible. Now an essentialist could consistently hold either position but of course not both. It is clear enough that essentialism is compatible with an opposition to strict determinism as outlined above. What may not be equally clear at first is that essentialism is also compatible with strict determinism. For consider the essentialist who is also a strict determinist. One

way in which they could expound their position is to invoke the notion of metaphysical possibility. On this view, while it is causally impossible that Napoleon won at Waterloo, it is metaphysically possible that he did so. That is to say that the laws of nature could have been sufficiently different to allow for Napoleon's victory; perhaps they were still strictly deterministic, perhaps not. The important thing for the essentialist is that they could have been different.

But in what sense of 'could' are we supposed to understand this claim: causal, logical or what? To say that something is causally possible is to say that its occurrence is consistent with the laws of nature. So with which laws of nature is the obtaining of different laws of nature consistent? To say that something is logically possible is to say that an assertion of its obtaining or occurring is not self-contradictory. But even an essentialist who opposes strict determinism can grant that it is not self-contradictory to assert that the laws of nature are other than they actually are. So when the essentialist-cum-strict-determinist says that the laws of nature could have been sufficiently different to allow for Napoleon's victory at Waterloo, they typically want to say that there is a third non-logical, non-causal sense of 'could', in which sense Napoleon could have won at Waterloo - in effect a metaphysical sense. Furthermore, even the essentialist opposed to strict determinism will want to acknowledge the existence of this third sense. Such an essentialist will want to say that Napoleon's victory is logically, causally and metaphysically possible.

But what does it mean to say that something is metaphysically possible? Again if we invoke the metaphor of a replay we could explicate this as follows: in a replay of the history of the universe things behave differently. Alternatively, as Kripke explains:

> Ordinarily when we ask intuitively whether something might have happened to a given object, we ask whether the universe could have gone on as it actually did up to a certain time, but diverged in its history from that time forward so that the vicissitudes of that object would have been different from that time forth. (Kripke, p.115.)

This notion of divergence would appear less metaphoric to commonsense than the notion of a replay; indeed some essentialists go on to postulate the existence of an infinitude of what they call 'divergent histories' or 'alternative universes or possible worlds' . In any case, for an essentialist there are some ways it is metaphysically possible for a thing to be, and other ways it is impossible for that thing to be. If it is metaphysically impossible for something not to be F, then being F is said to be essential to that thing - it is F in all the replays or alternative histories in which it figures - be it literally or figuratively - and so is essentially F. If the thing in question is F but metaphysically could have been not F, then it is accidently rather than essentially F - it is not F in some replay or alternative history. It is only on pain of circularity that I can explicate this notion of metaphysically-could-have without appealing to what I at any rate take to be at best **metaphors** of replays and divergent histories. I could of course appeal to the pre-analytic intuitions the majority of us brings to bear in its understanding of modal talk, but as I shall argue in chapter 2 these intuitions are incoherent. I will want to argue that we should not only eschew talk of a thing's having to be so but also resist the admittedly strong pre-analytic temptation to say that some things could have been otherwise. In short I am with Quine an anti-essentialist: I find the alleged distinction between essential and contingent properties incoherent, despite its admitted pre-analytic plausibility. Whether this is compatible with a necessitarian view of the causal relation is a question I shall not consider.

1.2.2 WOULD HAVE AND COULD HAVE

Anscombe has made remarks which suggest that she is also an anti-essentialist. She objects to the Molinist doctrine of *scientia media*, according to which

> God knew what anybody would have done if, e.g. he hadn't died when he did... [But] it seemed to me [Anscombe] that there was not... any such thing as what would have happened if what did happen had not happened, and... there was no such thing... as what someone would have done if ... (Anscombe 1981, p.vii.)

Anscombe is not just making the point that counterfactual would-have statements often suffer from an indeterminacy of truth-value, so that for example it is indeterminate whether it is true or false that Nixon would still have become President if Kennedy had not been shot. Even an essentialist could concede the indeterminacy of some counterfactuals. Her more basic point is that there are no such things as ways things could have been. This position strikes more deeply than a mere denial of a Lewisian realism about possible worlds: specifically, it is the thesis that it is simply incorrect or even unintelligible to talk about how things **could have** been. This is not to say that things **had** to be as they are, for on this view that is equally incorrect or unintelligible; all one can correctly or intelligibly say is that things just happen as they do. Indeed, I argue for this position in chapter 2. It follows that when one says to another that they shouldn't have done X, one is in effect saying something unintelligible if one thereby intends to suggest that the other **metaphysically** could have done otherwise; all one can legitimately intend to convey is that in future the other should bear in mind the consequences of doing X in certain circumstances. (I owe this point to John Mercier.) It follows that I reject a branching theory of time or the existence of non-actual alternative universes. (The nature and structure of time are questions too big to be addressed in this thesis.)

If talk about how things metaphysically could have been is illegitimate, what of talk

about how things metaphysically **could** be? That is to say, what are we to make of the statement that I **could** get up tomorrow one minute later than I actually will? I can make out only one coherent interpretation: **for all I know now**, I will get up tomorrow at a certain time, which time as it turns out tomorrow is one minute later than the time at which I actually do get up. (See chapter 3, section 3.2 for a defence of this claim.) This is the 'could' of **epistemic** possibility: my current **knowledge** is consistent with my getting up later tomorrow. And if the universe is indeterministic, then it might be that **no** one yet knows that I will not get up later, i.e. it is consistent with **everyone's** knowledge that I get up later tomorrow. (Cf. Aristotle on the future sea battle.)

Now we use and understand counterfactual talk all the time, which is why indeed we have a subjunctive. With such talk we assert logical or causal connections between circumstances expressed in the antecedent and those expressed in the consequent of the counterfactual. For example I might say that if you were my parent then I would be your child, or that if you had not eaten those mushrooms, you would not have had that stomach ache. I do not mean to say either that you actually were my parent or even that you could have been. Rather what I intend to do is to draw out the conceptual connections between being a parent and being a child. Similarly I do not mean to say that you did not eat the mushrooms or even that you could have desisted from eating them. My intention is rather to bring out what I take to be certain causal connections between the ingestion of certain toxins and certain physiological processes. So would-have statements are perfectly intelligible. They are intelligible even when the antecedent when literally understood expresses a logically impossible state of affairs. Thus the antecedent of the statement that if I had been Kennedy then I would have sacked Hoover, literally understood, makes reference to a self - contradictory state of affairs, namely my being someone distinct

from me; I cannot after all both be I and not be I. Nevertheless, it is clear what I intend with such a statement: I mean neither that I was Kennedy nor that I could have been Kennedy, but rather that if certain circumstances confronted me then I would decide on a certain course of action other than the one followed by Kennedy in those circumstances, and that it would have been better for Kennedy had he acted similarly, even though he **metaphysically** could not have acted differently from the way he actually did act. The upshot is that one can endorse the intelligibility of counterfactual talk without thereby committing oneself to the intelligibility of metaphysically-could-have talk.So if the anti-essentialist meant by the claim that a could have been F **merely** that, had circumstances been such and such, a would have been F, there would be no real problem of understanding this claim. But typically the full-essentialist means more than this when they make such assertions. They typically mean that alternative non-actual outcomes **metaphysically** could have happended. And such assertions I shall argue in chapter 2 lead to incoherence.

It might be replied that I have overlooked a simple argument for the coherence of metaphysical possibility: surely I could get up tomorrow one minute later than I actually will. So wouldn't it be true to say **after** I get up tomorrow that I **could have** got up one minute later than I actually did? I would respond that this argument commits the fallacy of equivocation: it trades on an ambiguity in the term 'could'. If the argument is not to be a *petitio*, then the sense of 'could' in the first sentence can't be that of 'metaphysically could '. So the only other way to understand it is in terms of **epistemic** possibility, i.e. we must read the first sentence as saying that **for all I know**, I will get up tomorrow at a certain time - which time as it turns out **tomorrow** is one minute after the time I do in fact get up at. But since the **conclusion** requires the metaphysical sense of 'could', it follows that the argument commits the fallacy of equivocation.

1.2.3 METAPHYSICAL ESSENTIALISM VERSUS CONCEPTUAL ESSENTIALISM

At this point the essentialist might object that they can explain the import of essentialist claims without recourse to the idiom of metaphysical possibility and its attendant metaphysic of replays or alternative histories or worlds. On this view, **counterfactual would-have** statements as opposed to **metaphysically could-have** statements are all that is required to express the distinction between contingent and necessary properties. For example, the essentialist could urge that all they mean in saying that Napoleon could have won at Waterloo is that the **concept** of Napoleon does not include his defeat at Waterloo. To say Napoleon won at Waterloo is to say something false but consistent with the **concept** of Napoleon. Expressed as a **would-have** statement, it might read: 'if Napoleon had won at Waterloo, it would indeed have been Napoleon who won and not someone else'. But there is no suggestion that Napoleon **metaphysically could have** won, i.e. that he appears as a victor in some playback or alternative history. All that is being said is that the concept of Napoleon does not include his defeat at Waterloo. But nor does it include his winning at Waterloo, for if it did, then it wouldn't ever be true that Napoleon lost at Waterloo - not at least if the concept of Napoleon is a **true** concept of him.

Similarly, the essentialist could urge that all they mean in saying that Napoleon could not have been a clay-pipe is that the concept of Napoleon does include his being something which is inconsistent with being a clay-pipe, viz. his. being a human being. To say Napoleon was a clay-pipe, is to say something both false and inconsistent with the concept of Napoleon. Expressed in terms of a would-have statement, it would read: 'if Napoleon had been a clay-pipe, then it wouldn't have been Napoleon who was the clay-pipe' or 'if anything had been a clay-pipe, then it

wouldn't have been Napoleon'. But there is no suggestion here that Napoleon **metaphysically** had to be a non-pipe, i.e. that he appears as a non-pipe in each of the entirety of replays or alternative histories. All that is being said is that the **concept** of Napoleon includes his being a non-pipe.

The upshot, so the essentialist can conclude, is that the distinction between contingent and necessary propeties **can** be expressed without the metaphysical apparatus of replays or alternative histories or universes; essentialism does not require the notions of **metaphysical** possibility and impossibility. Instead it manages perfectly well with an appeal to the notions of **conceptual** possibility and impossibility. On this view there are such things as **concepts** of objects: it then follows that a contingent property is one that is **conceptually possible**, i.e. it is not **included** in the concept of the object, but nor is its contradictory included in the concept. An essential property for its part is one that is **conceptually necessary**, i.e. it **is** included in the concept of the object. Compare on this point a discussion by Arnauld in his correspondence with Leibniz:

> I must consider as contained in the individual concept of myself only that which is such that I should no longer be me if it were not in me:...all...[which] is...such that it could be or not be in me without my ceasing to be me, cannot be considered as being contained in my individual concept. (Mason, p.30)

According to the essentialist, the advantage for them of this explication of essentialism in terms of **conceptual** possibility and necessity rather than metaphysical possibility and necessity is that it avoids the objections which I shall raise in chapter 2 against the notions of metaphysical possibility and necessity, which notions I shall argue lead to incoherence. Furthermore, on this view, all necessity is analyticity, for a necessary *de re* statement is necessary in virtue of the logical entailment of the predicate by the subject term designating the concept of the

statement's subject. I call this general position 'conceptual essentialism', as opposed to metaphysical essentialism.

Now my response to this position is to counter that objects do not have concepts, or at least that if they do, any concept comprising less than **all** of an object's sundry properties is **conventional**; so that any distinction between contingent and necessary properties is not grounded *de re*, i.e. not grounded in the thing itself, but is itself a matter of stipulation and thus arbitrary and hence of no metaphysical significance. For I hope to persuade the reader in the course of this thesis that **every** property of an object is on an equal **metaphysical** if not practical footing.

If my attack on the coherency of metaphysical essentialism proves to be well-founded, then one consequence will be that teachers of introductory courses in logic should desist from explaining the notion of a valid argument 'p hence q' in terms of 'the **intuitive** [my emphasis] notion of implication, which requires, not just that 'p'is not true and 'q' false, but that 'p' **could** [i.e. **metaphysically** could] not be true and 'q' false.' (Haack 1978, p.176.) It is a commonplace that it was the failure - as perceived by C.I.Lewis - of *Principia*'s material implication to do justice to this intuitive sense of 'implication' - to my mind probably understood by him implicitly if not explicitly in terms of **metaphysical** possibility - that motivated Lewis' development of modal logic in the first place. At any rate that the relevant sense of 'could' intended by **teachers** of logic is that of '**metaphysically** could', becomes apparent once we note how they typically go on to explain to their students that in contrast with the case of a **valid** argument with true premises, in the case of an **invalid** argument with true premises and true conclusion, while the circumstances described by 'q' may have actually obtained they **need** not have, i.e. things **metaphysically** could have been otherwise, even if the circumstances described by 'p' had still obtained. Compare Sprigge on this point:

> To say that...[a proposition] is synthetic suggests the thought that ... it might have been otherwise ... If we say that this proposition ['it is a building', where 'it' refers to the National Gallery] is synthetic it suggests that it is ... conceivable... that the National Gallery should not have been a building. Then we worry ourselves as to what it would have been like for the National Gallery not to have been a building. For if it is a synthetic proposition that it is a building, it must be logically possible ... though false ... that it is not a building...[But then] we should be able to imagine what it would have been like for ...[this] to be true [which however we have difficulty in doing]....[T]o call ...[singular] propositions synthetic is misleading, in that it suggests ... that one could imagine what it would be like for them to be false. ... Here is a certain thing. ... Certain descriptions apply to the thing. Once one has noticed that they apply, one does not always know what to make of the question - must they have applied? To ask whether a world is conceivable in which this very thing was quite otherwise is bewildering....Thus it would be in some ways more suitable to say that the analytic-synthetic division does not apply to ...[singular] propositions than to say that they are synthetic. They cannot properly be called analytic [if the subject term functions purely to refer]. On the other hand one cannot - in many cases - properly ask what it would be like for them to be false. If, therefore, we do call them synthetic, it should only be as a way of contradicting the claim that any of them are analytic. (Sprigge 1962, p.207.)

What is interesting from my anti-essentialist point of view is, firstly, Sprigge's thought that it is illegitimate to inquire of an object whether it could have been otherwise in a **metaphysical** sense of 'could' (he does allow limited talk of **causally** possible alternatives, though he does not seem to appreciate the distinction), and, secondly, the tendency he highlights among logicians and philosophers to appeal to talk of things' having been otherwise when explicating non-analytic propositions and thus conditionalisations of invalid arguments. Sprigge's observations remain good even when stripped of their dependence on the dubious assumption that the criterion of the logically possible is conceivability or imaginability.

In the remainder of this chapter I address the subject of **metaphysical** essentialism as opposed to **conceptual** essentialism. Henceforth all references to essentialism in the remainder of this chapter are to be understood as references to metaphysical essentialism, unless expressly indicated otherwise.

1.2.4 FORMAL AND ORDINARY LANGUAGE CHARACTERISATIONS OF ESSENTIALISM

It is important to emphasise that when I say the full-essentialist thinks that some of a particular object's (if not necessarily every object's) properties are essential to that object, I mean that the object in question is held to possess its essential properties **absolutely**, i.e. independently of its mode of designation or specification or description. Quine would grant - so far as he would make any concession to necessity - that an object could have some property necessarily, but only under some specification or description. Essential properties are for Quine **relativised** to specification: in his view, just as it is meaningless to talk of some object's being tall absolutely, so it is meaningless to talk of some object's being necessarily so and so absolutely. Hence one could call Quine 'a conceptual essentialist'; but of course given his doubts about analyticity, he might resist his being so characterised.

Individualist essentialists could agree with catholic essentialists that (A): some properties, such as the property of self-identity (Wittgensteinian reservations aside), are essential to every object. The catholic essentialist cannot however allow possibilities open to the individualist: viz. that (B): some properties, while not possessed by every object, are possessed **essentially** by every object that does possess them; or that (C): some properties are possessed **essentially** only by some but not all of those objects which do in fact possess them. Let us call any essentialism endorsing (A) '**A WEAK ESSENTIALISM**'. Any full-essentialism endorsing either (B) or (C) will be called '**A STRONG ESSENTIALISM**'. '**LEIBNIZIAN ESSENTIALISM**' I define as the essentialist theory which holds that every one of any object's properties is essential to that object. The following tabulation of my essentialist taxonomy indicates the various compossibilities obtaining among the different forms of full-essentialism described.

	Catholic	Individualist	Leibnizian	Strong
Weak	+	+	+	+
Strong	-	+	+	
Leibnizian	-	+		
Individualist	-			

where

	j
i	+

indicates that i and j are compossible

and

	j
i	-

indicates that i and j are not compossible.

Quine's characterisation of Aristotelian essentialism quoted at the bottom of p.1 renders it compatible with all the forms tabulated save Leibnizian essentialism. While weak essentialism is compatible with all the other forms tabulated, it is entailed only by catholic essentialism.

At this stage there arises an important question: does Quine's modal characterisation of Aristotelian essentialism square with his ordinary language characterisation of it? This question raises the important problem of a **semantics** for systems of QML, which must be addressed now, all the more so since Quine's strictures against systems of QML apply to them as **interpreted** systems, i.e. his objections arise primarily at the level of semantics rather than that of syntax. Indeed, the syntactic misgivings he does express are motivated by semantic considerations. Quine would not deny that an uninterpreted formal system could be contructed whose vocabulary and formation rules would allow as well-formed such formulae as

'necFa', 'nec(Ex)Fx' and '(Ex)necFx' (using 'nec' in place of the logician's box). Various axioms and rules of inference could than be stipulated for the different systems of QML, which would allow, for example, the proof of 'Fav~Fa |- nec(Fav~Fa)'. Presumably we could also endorse the proof of 'necFx |- (Ex)necFx' or '(x)necFx |- necFa'. Provided of course that these **uninterpreted** formal systems are consistent, they would be harmless enough. So since the issue of a semantics for systems of QML is all-important, I must now digress on matters semantical; among the results of my digression will be a determination of how to understand Quine's formal (canonical) characterisation of Aristotelian essentialism and of whether it squares with his ordinary language characterisation.

1.3 INFORMAL MODAL SEMANTICS

1.3.1 NECESSARILY TRUE THAT P

It is enough for my purposes that the exposition of the semantics be non-technical; I will explain in terms of ordinary language just which features of ordinary language the modal vocabulary of systems of QML is intended to capture.

Clearly, the notion of alethic necessity is rooted in commonsense; moreover, one's pre-analytic understanding of this notion is typically considered to be intuitive: letting 'p' represent any proposition, we commonly take ourselves to understand more or less what someone means when they say that it is necessarily the case that p; or that it is necessarily true that p; or that p is necessarily true or necessarily the case; or that necessarily p; or that p is necessary; or - at a greater remove from idiomatic usage- that the state of affairs described by the proposition that p is necessary. For the essentialist alethic necessity involves not only the necessity of logical necessity but also the necessity of metaphysical necessity, as opposed to the necessity of epistemic necessity or deontic necessity or causal necessity. This is not

to say that for the essentialist the same notion of necessity may not be operative throughout, rather that their and hence my concern is merely with alethic necessity. This immediately invites an important question: are the notions of logical necessity and metaphysical necessity identical, or are there, as the essentialist would urge, grounds for a distinction?

To say that it is necessarily the case that p is to say that p is necessarily true. Now an important way in which propositions are standardly thought to be necessarily true is for them to be (a) logically necessary (or logically true or valid). Such propositions are necessarily true in virtue of their logical form alone. Other propositions then count as necessarily true in virtue of (b) a combination of their logical form and the meaning of the constituent terms of their corresponding sentences. Such propositions are reducible to logically necessary propositions by the substitution of synonyms. In both cases (a) and (b) the denial of such propositions is self-contradictory, and thus deemed to be necessarily false. Propositions falling in these categories are said to be analytic. This understanding of analyticity derives of course from Frege. It is commonly said, not just by avowed essentialists, that analytic propositions **must** be true - for they are necessarily true and thus necessary truths. Now the empiricist who turns out typically to be anti-essentialist would have it that the class of analytic propositions exhausts the class of necessary truths, but at any rate the **full-essentialist** would appear to have to accept the existence of some non-analytic (synthetic) necessary truths. To see this consider the proposition concerning that paradigmatic individual Socrates that he (Socrates) is not an amoeba. It is safe to assume that a **full- essentialist** would typically deem this proposition to be necessarily true, notwithstanding our being hard pressed to render this as **analytically** true. Of course attempts have been made to force apparently synthetic necessary truths into an analytic mould. But for the present,

lest I be accused by the essentialist of prejudging the issue of the possibility of the existence of synthetic necessary truths, I shall recognise the category of metaphysically necessary truths as a broader category of truths than that of analytic truths which it includes. For their part, the essentialist might urge that this recognition of the broader category of metaphysically necessary truths does not necessarily introduce a new kind of **necessity** over and above that operating in logical necessity; for, so they might urge, if one accepts that the proposition that Socrates is not an amoeba is necessarily true, then surely in the same sense of 'must' in which 'pv~p' **must** be true, or in which 'all bachelors are unmarried' **must** be true, so 'Socrates is not an amoeba' **must** be true. The relevant logical difference - if there is one - between the first two propositions and the third will reside in **that** in virtue of which they must be true. We shall henceforth understand talk of necessary truths as talk about metaphysically necessary truths, leaving open for the time being the question whether there are in reality necessary truths which are metaphysically but not analytically necessary. That there are in fact none is to be anticipated from my rejection as incoherent of the notion of metaphysical possibility in chapter 2.

1.3.2 NECESSARILY-IS F

A necessary truth is necessarily true, i.e. it must be true; it cannot not be true. (This characterisation of what are commonly called 'necessary truths' will have to abandoned in the light of difficulties raised in chapter 2 for the notion of metaphysical necessity.) Now the essentialist could urge that if one grants that it is necessarily true **that** Socrates is not an amoeba, could one not then equally grant that it is necessarily true **of** Socrates that he is not an amoeba? But what would this mean? We know what it is for something to be true **of** a subject as opposed to being **necessarily** true of it. For example, if it is true **that** Socrates is (was) a philosopher, then it is true **of** Socrates that he is a philosopher. Conversely, if it is

true **of** Socrates that he is a philosopher, then it is true of Socrates **that** Socrates is a philosopher. Indeed, Socrates **truly-is** a philosopher.

By parity of reasoning then, the essentialist might urge, if it is **necessarily** true (it must be true; it cannot not be true) that Socrates is not an amoeba, then it is necessarily true **of** Socrates that he is not an amoeba. That is to say, it cannot not be true of Socrates that he is (such that he is) not an amoeba. Conversely, if it is necessarily true **of** Socrates that he is not an amoeba, then it is necessarily true **that** he is not an amoeba. Let us say that a name is used referentially or has referential force just where its semantic function is to refer; its sense, if it has one, is semantically idle. Where 'a' is a non-empty name with **referential force**, let us refer to the thesis that (A) it is necessarily true that a is F if and only if (B) it is necessarily true of a that it is F as 'the parity thesis'.[1] The essentialist would typically endorse the parity thesis: they would endorse not only that B only if A but also A only if B. To an **anti-essentialist** this thesis already seems to concede too much, since **being necessarily true of Socrates** (say) seems to abandon the grounding of necessity in meaning - for 'a' is being used with referential force - and thus to bring in a ***de re*** conception of necessity. And of course the worrying question for the anti-essentialist is just what it could **mean** to go beyond the *de dicto*. For the anti-essentialist would be willing to allow that some sentences of the form 'a is F' are necessarily true ***de dicto***, provided of course one subscribes to a non-Millian view of names. (For a Millian anti-essentialist no sentence of the form 'a is F' is necessarily

---

1 The parity thesis captures an aspect of the distinction between *de dicto* and *de re* readings of a sentence. In the former case the **proposition** is said to bear a property essentially, viz. the property of being true, while in the latter case it is an extra-linguistic and non-propositional individual that is said to bear a property essentially.

true *de dicto*.) But while the possibility of *de dicto* necessity does arise for the non-Millian anti-essentialist, they reject on conventionalist grounds the transition to ***de re*** necessity. For the present, I ask the indulgence of the anti-essentialist. I hereby stress that it is in accordance with the methodological principle of charity that I let the parity thesis stand. This is a provisional manoeuvre only, executed in order not to prejudge at the very beginning - by exclusion of the possibility of a *de re* conception of necessity - the issues of essentialism and full-essentialism. Anti-essentialist worries about *de re* necessity will be duly addressed and vindicated.

Now if it is admissible to speak of its being necessarily true of Socrates that he is not an amoeba, can we then go on to speak of its being true of Socrates that he necessarily-is not an amoeba? But what would it mean to say that Socrates necessarily-is an F, where F is either a positive or negative predicate? From a grammatical point of view, 'necessarily' as it occurs in the phrase 'it is necessarily true that p' occurs as an adverb modifying the adjective 'true', which in turn modifies the clause 'that p'. The state of affairs expressed by this clause is said to obtain, and moreover to obtain of necessity, i.e. it necessarily obtains. Now notice that 'necessarily', as it occurs in the last clause of the previous sentence, viz. 'it necessarily obtains', modifies a verb. Since from a logical as opposed to a grammatical point of view a copula-cum-adjective construction or a copula-cum-substantive construction is on a par with a verb-construction, in as much as both function predicatively, it would seem that 'necessarily' in both mentioned occurrences could be construed as qualifying the predicate. In particular, 'necessarily' could be interpreted as functioning to characterise the predicate's **mode** of characterising the subject - hence the appellation of 'alethic **modality**'; thus: 'is' ('truly-is' / 'actually-is' / 'really-is'); 'necessarily-is'; and 'possibly-is'.

Of course in classical logic, i.e. bivalent truth-functional logic in the tradition of Russell and Quine, 'necessarily' is standardly treated as an unary **sentence** operator; it operates on a sentence (open or closed)[2] .It certainly does not function as an operator on predicates to form one predicate from another. This parallels the treatment of the tilde in classical logic as an unary sentence operator,where standardly tilde operates on sentences, not on predicates. Nonetheless some logicians including Geach have advocated just such a construction, calling it 'internal negation', as opposed to so-called external negation in which tilde governs a sentence. Intuitively,'(~F)a' is true just where a lacks the property of Fness.

So there is a precedent in non-modal logic for attaching an unary sentence operator to a predicate. Could the essentialist appeal to this precedent with respect to 'necessarily' in modal logic? To do so, they would have to secure a well-defined meaning for 'necF'. Providing truth-conditions for the contexts in which it appears

---

2 Quine needless to say has no truck with modal propositional logic, let alone QML. Nor did Russell, who disdained it in as much as: (1) it was irrelevant to one of his prime logical concerns, viz. the formalisation of a language adequate for foundational studies in mathematics, and in as much furthermore as: (2) modal notions resisted reduction as required by his program of logical atomism (see Rescher, pp.85-96). It is interesting in this respect to compare Quine, who thinks that tne notion of necessity would be acceptable, even if it were irreducible, if it were useful in integrating or simpifying our system of the world, which he thinks it isn't (see Quine 1990, p.244). In particular he thinks the notion of necessity unnecessary to explain the counterfactuals which are underwritten by laws of nature.

would go a long way in achieving this. Suppose then, given the essentialist's pre-analytic intuitions about possibility - which I shall argue in chapter 2 are incoherent - that '(necF)a' is to be true on an interpretation I just when the referent I(a) of the individual constant 'a' is such that it is necessarily true of it that it is F. Call this supposition 'the semantic precedent'. It follows that for all **non-empty** names 'a' **with referential force**, '(necF)a' is logically equivalent to 'nec(Fa)', for if '(necF)a' is true on I just when I(a) is such that it is necessarily true of I(a) that it is F, then given the parity thesis, it is necessarily true that I(a) is F, i.e. 'nec(Fa)' is true. Conversely, if 'nec(Fa)' is true, then it is necessarily true of I(a) that it is F, and thus true that (necF)a. So if Socrates **necessarily-is** a non-amoeba, then Socrates **must** be such that he is not an amoeba; he cannot not be such as not to be an amoeba. (It follows that Socrates must be a non- amoeba; he cannot be an amoeba.)

The *ad hominem* objection might be raised against this analysis of '(necF)a' that it renders a a necessary being, for not even full-essentialists would typically want to endorse the necessary existence of everything. For if '(necF)a' is true just where it is necessarily true of a that it is F, then if a did not exist, it would not be F (nor would it be not-F either) and so it would be false of a that it is F. Hence it would not be necessarily true of a that it is F, and thus '(necF)a' would be false. The truth of '(necF)a' requires the truth of '(necExist)a'. But this argument is fallacious: if 'Fa' entails that a exists, then it entails only that a exists; it does not entail that a **necessarily** exists; so if 'Fa' entails only that a exists, similarly '(necF)a' entails only that a exists, not that a **necessarily** exists. To think otherwise is to confuse the necessity of the entailment with the necessity of what is entailed..

Alternatively, it might be objected as an *ad hominem* against the essentialist who might subscribe to this semantics that, given that 'nec(Fa)' is true for non-empty 'a'

with referential force just where '(necF)a' is true, 'nec(Fa)' **entails** that a exists, since '(necF)a' does. But the essentialist would then have the case *per impossibile* where a necessary truth (viz. 'Fa') entails a contingent truth (viz. 'a exists') - on the supposition of course that not even full-essentialists would typically ascribe necessary existence to garden-variety concrete individuals. After all, it is an elementary point of logic that necessary truths entail **necessary** truths only. So don't the semantics I have introduced then commit the fallacy of what Chisholm calls 'inferring *necessitas de dicto* from *necessitas de re*' (Chisholm,p.84)? The essentialist must concede the force of this objection. Their only way out is to deny that 'Fa' or '(necF)a' **entails** that a exists. In support of this manoeuvre they might observe that it makes it possible for them to entertain talk of the essential properties of **non-actual** individuals (admittedly small consolation for the anti-essentialist). In any case, there is something to be said for the free logician's use of a non-existential partitive quantifier; this then becomes consistent with a substitutional approach to modal quantification to be introduced forthwith.

1.3.2.1 (∃x)(necF)x and (∃x) nec (Fx)

Given '(necF)a', the essentialist could decide to go ahead and generalise existentially on 'a', obtaining '(∃x)(necF)x'. A reading of this by their lights as 'something is such that it is true of it that it necessarily-is F' would make it true just where there is a true **substitution instance** of the form '(necF)a'. From this reading they could infer by the semantic precedent introduced in the previous section that something is such that it is necessarily true of it that it is F. And from this they could infer by the parity thesis that something is such that it is necessarily true that it is F. But this would naturally enough constitute, at least for the essentialist, a reading of '(∃x)(nec(Fx))', which could be inferred from 'nec(Fa)' by the following rule of existential generalisation: '(∃x)nec(Fx)' would be true just where it had an

appropriate (see the next paragraph) substitution instance of the form 'nec(Fa)'. And given '(∃x)nec(Fx)', something is such that it is necessarily true that it is F, the essentialist could go on to infer by the parity thesis that something is such that it is necessarily true of it that it is F, and thus infer by the semantic precedent that something necessarily-is F. But this would constitute a natural reading for the essentialist of '(∃x)(necF)x'. So '(∃x)(necF)x' and '(∃x)(nec(Fx))' would be interderivable and thus logically equivalent.

This line of reasoning commits QML to substitutional quantification: '(∃x)(necF)x' is true given some true substitution instance '(necF)a', and '(∃x)nec(Fx)' is true given some true ***de re*** substitution instance 'nec(Fa)'. The full-essentialist could not let just any true ***de dicto*** substitution instance suffice for the truth of '(∃x)nec(Fx)'. For consider 'The teacher of Plato taught Plato', where 'F' = 'taught Plato' and 'a' = 'the teacher of Plato'. The proposition 'Fa' could be conceded by the non-Millian **anti**-essentialist to be necessarily true but only *de dicto*, not *de re*. To that extent they would allow 'nec(Fa)'. Similarly the **full**-essentialist would readily allow a *de dicto* reading of 'nec(Fa)', but disallow a *de re* reading, for typically the full-essentialist would hardly want to count Socrates' tutorship of Plato as essential to him. The full-essentialist would typically want to say that **Socrates** might not have tutored Plato and yet still have been **Socrates.** So some true *de dicto* substitution instances will be discounted by the full-essentialist when determining the truth-value of '(∃x)nec(Fx)'. At this point the anti-essentialist will justifiably demand to know the basis on which such discrimination between substitution instances is to be made. They will furthermore enquire after the basis on which in general sentences of the form '(necF)a' are determined to be true or false. At this point the full-essentialist will reply that the basis is none other than a theory of essential properties. It will be my aim to examine some such theories in the following chapters, in order to

establish whether the full-essentialist can vindicate their reply to the anti-essentialist.

1.3.2.2 (x)(necF)x and (x)nec(Fx)

Interpretations of '(x)(necF)x' and '(x)nec(Fx)' are forthcoming. The essentialist could urge a reading of the former as 'everything is such that it is necessarily is F'. It is true just where every substitution instance '(necF)a' is true. The essentialist could read the latter '(x)nec(Fx)' as everything is such that it is necessarily true that it is F. This formula will be true just where every **appropriate** substitution instance 'nec(Fa)' is true (see 1.3.2.1).

1.3.3 EQUIVALENCE OF '(necF)a' AND 'nec(Fa)'

1.3.3.1 EMPTY NAMES 'A'

The parity thesis and the semantic precedent were introduced with non-empty names in mind. For such names, '(necF)a' and 'nec(Fa)' are logically equivalent. Whether '(necF)a' **in general** is logically equivalent to 'nec(Fa)' depends on the stipulations we make concerning the truth-conditions of predications 'Fa'involving an empty name 'a': is 'Fa' to be false (or indeterminate) *tout court*, or can it be true or false depending on the domains involved ? The truth-table below tabulates the various possibilities for predications involving an empty name 'a'.

	T or F	F	Indeterminate
Fa	T	F	I
(necF)a	T F	F	I
nec(Fa)	T F	F	I

Under rubric 'T or F', 'Fa' is true just where a non-existent F is allowed to

correspond to 'a' and is false otherwise. If 'Fa' is true, then '(necF)a' is either true or false. In either case, '(necF)a' will be materially equivalent with 'nec(Fa)'. If 'Fa' is false, then '(necF)a' and 'nec(Fa)' will both be false. So on this semantics '(necF)a' is logically equivalent to 'nec(Fa)' for all 'a', empty and non-empty.

Under rubric 'F', 'Fa' entails that a exists. So if 'a' is empty, 'Fa' is false. Hence '(necF)a' and 'nec(Fa)' are false. Again '(necF)a' is logically equivalent to 'nec(Fa)' for all 'a' on this semantics.

Under rubric 'Indeterminate', 'Fa' is indeterminate if a does not exist. Hence '(necF)a' will also be indeterminate. But since it is indeterminate and thus not **true** that Fa, 'nec(Fa)' will be **false**. Thus '(necF)a' will not in general be logically equivalent to 'nec(Fa)' on this Strawsonian semantics.

1.3.3.2 RUSSELLIAN INTERPRETATION OF DEFINITE DESCRIPTIONS

Leaving these observations concerning names, we note that Quine in any case will not countenance singular terms. Whether proper names or definite descriptions, they are to be eliminated in canonical paraphrase via Russell's theory of definite descriptions, supplemented by Quine's predicative parsing of proper names. This in turn will bear on the issue of the equivalence of '(necF)a' and 'nec(Fa)'. So if we take a Russellian approach to grammatically singular descriptions, external 'nec(G((ix)Fx))' translates into (1) 'nec($\exists$x)((Fx & (y)(Fy$\rightarrow$y = x)) & Gx)', while internal '(necG)((ix)Fx)' translates into (2) '($\exists$x)((Fx & (y)(Fy$\rightarrow$ y = x)) & (necG)x)'.

On essentialist intuitions (1) reads as 'it is necessarily true that some object is both the only F and also is G'. The latter formula (2) can be understood as saying that some object is such that it is the only F and such that it necessarily is G. But even

allowing that '(ix)Fx' has a denotatum, i.e. that there exists an item which uniquely satisfies the description, the former translation (1) is not equivalent to the latter (2), for in (1) all that is stated is that there must be something or other uniquely F which is G, but **not** that whatever turns out to fill the bill is such that necessarily **it** is G. To illustrate this more concretely: letting 'F' represent 'is the number of planets in the earth's solar system' and 'G' is odd', 'nec(∃x)((Fx & (y)(Fy → y=x)) & Gx)' will come out false, for the full-essentialist would typically urge that there appears to be nothing necessary (necessarily true) in there being a solar system in the first place, and so nothing necessary in there being a true predication of unique Fness of **any** number, let alone the one (viz. nine) that actually fills the bill. But there being such a predication is a necessary condition of the truth of 'nec((Ex) ...' . Now contrast the falsity of 'nec((∃x) ...' with the truth of '(∃x)((Fx & (y)(Fy → y=x) & (nec G)x)' on the same interpretation. The latter formula comes out as true just because it has a true substitution instance, viz. '((Fa & (y)(Fy → y = a)) & nec(Ga))', where 'a' is '9'. So on a Russellian interpretation of 'G(ix)Fx', 'nec(G((ix)Fx))' will not in general be equivalent to '(necG)((ix)Fx)'.

1.3.3.3 A PARTICULAR NON-RUSSELLIAN INTERPRETATION OF DEFINITE DESCRIPTIONS

Turning now from a Russellian to a non-Russellian approach to definite descriptions, where we count definite descriptions as genuine **singular** terms, there arises the question of what to make of their function as singular terms in contexts both external 'nec(G((ix)Fx))' and internal '(necG)((ix)Fx)', for in both schemata there hides an ambiguity. On the one hand, '(ix)Fx' could be functioning as what Kripke calls 'a rigid designator', i.e. when the semantics governing the intended interpretation are invoked, the denotation of '(ix)Fx' remains invariant in all contexts. On the other hand, '(ix)Fx' could also be functioning as what Ravnkilde,

inspired by Donnellan's seminal 'Reference and Definite Descriptions' of 1966, calls 'the universally attributive use of a definite description'. Following Donnellan, Ravnkilde distinguishes between the referential use of definite descriptions and their attributive use. The nomenclature is unfortunate - as Donnellan himself concedes in later work - in as much as it suggests that attributive uses of definite descriptions lack a referring function, which of course is contrary to the intention of Donnellan's distinction. Donnellan did not intend the distinction as one between referring and non-referring uses of definite descriptions. Rather by 'referential' Donnellan intends the situation where in asserting 'G((ix)Fx)' one singles out an object in order to make it the subject of a certain predication, but what one **asserts** is merely that a particular object x is G, so that a misdescription (whether intentional or not) of x as '(ix)Fx' does not thereby render the assertion false - for that misdescription is not part of what is asserted of the subject; what is asserted of it is that it is G, not that it is F as well. The subject's being G is sufficient to render 'G((ix)(Fx)' true. As Ravnkilde explains:

> when a definite description is used referentially, its sense yields no conclusive clue to the identity of the intended referent. The intended referent and the denotatum ... must be kept distinct. They may indeed coincide, but then again they may not. (Ravnkilde, p.14)

The speaker who singles out x referentially as '(ix)Fx' for the purpose of predication will have the means to identify x independently of the referential description, drawing on descriptions based on personal acquaintance, even if the acquaintance is so slight as to yield merely the impoverished description 'the person whose face comes to my mind as I now single him out to you by the use of '(ix)Fx''. The use of this description is solely to direct the hearer's attention to the speaker's intended subject of discussion and may succeed even if it actually misdescribes the intended referent.

Now contrast this referential use of definite descriptions with the **attributive** use of definite descriptions. In this case the speaker might have no personal acquaintance at all with the intended subject of discussion. In this situation the referent and denotatum **must** coincide. The speaker means to refer to whoever or whatever satisfies the description. In asserting that the F is G, the speaker asserts **both** that the subject is F **and** that it is G, so the assertion is falsified by a misdescription of the subject as F. Ravnkilde points out that the speaker might not be in any privileged position as against the hearer in the matter of determining the referent (Ravnkilde, p.15). The speaker might not be able to identify the denotatum independently of its being the object satisfying the attributive description (Ravnkilde, p.14).

Adopting Donnellan's distinction between the referential and the attributive, Ravnkilde goes on to extend it, dividing attributive uses into two kinds: what he calls 'singular attributive' and 'universal attributive'. The singular attributive use coincides with Donnellan's attributive use; in contrast, the universal attributive use does not pick out an object of the lowest order, i.e. those things - individuals - which can be subjects of predication but cannot intelligibly be said to be themselves predicable. Rather, this use picks out higher order objects such as a particular class (or corresponding property), of which it is then asserted that membership of the class (or instantiation of the property) is sufficient for membership of another class (or instantiation of another property) (see Ravnkilde, pp.15-16). For an illustration consider the equivalence between 'the whale is a mammal' and 'all whales are mammals', i.e. every thing that is a whale is also a mammal. In what follows we confine ourselves to the above non-Russellian account of definte descriptions and disregard other accounts such as Frege's, Strawson's, Hilbert's, etc.

1.3.3.4 INTERPRETING 'necG((ix)Fx)', WHERE '(ix)Fx' IS CONSTRUED AS A SINGULAR TERM

We shall now gather together these strands in order to determine how they bear on the interpretation of formulae such as (1) 'nec(G((ix)Fx))' and (2) '(necG)((ix)Fx)', where '(ix)Fx' is understood as functioning as a singular term.

(1): where '(ix)Fx' is used referentially and the denotatum exists, then our earlier observations about non-empty proper names (individual constants) in the fifth paragraph of 1.3.2 apply. If '(ix)Fx' is interpreted as functioning singular attributively and the denotatum exists, then (1) will be true just where it is necessarily true of a certain object x that it is both the unique F and G, i.e. that it is necessarily true that the object x which is the unique F is also G. If '(ix)Fx' is interpreted as functioning universal attributively, then (1) will be true just where it is necessarily true of **every** F that it is G, i.e. it is necessarily true that every F is G. Now clearly a singular attributive interpretation of a sentence may diverge in truth-value from a univeral attributive interpretation of a sentence may diverge in truth-value from a universal attributive interpretation of the same sentence. Thus it might be true of a certain presently existing object that happens to be uniquely F that it is G, without its being true of past or future Fs that they were or will be G, hence without its being necessarily true that all Fs are Gs.

(2): on a singular attributive interpretation of '(ix)Fx', (2) will come out true on essentialist intuitions just where the object which is uniquely F is necessarily G, i.e. just where it is necessarily true that this object is G (although this doesn't mean much to the anti-essentialist). On a universal attributive interpretation, (2) will be true on essentialist intuitions just where it is true of every F that it is necessarily G, i.e. just where it is necessarily true that every F is G. Again, it is easily seen that the two interpretations of the same sentence may diverge in truth-value.

Note that (1) and (2) are logically equivalent on an interpretation of '(ix)Fx' as singular attributive, and that similarly (1) and (2) are logically equivalent on a universal attributive interpretation, for on the former interpretation (1) and (2) have the same truth-conditions and both do in fact satisfy them, and similarly on the latter interpretation (1) and (2) have identical truth-conditions, and both do in fact satisfy these.

1.3.4 CONTINGENT TRUTHS MAY BE UNDERSTOOD AS NECESSARY, AND CONVERSELY

Two important results emerge from our discussion. They are that certain propositions commonly held by the anti-essentialist to be **not** necessarily true ***de re*** may be given a reading on which they **are** thought to be necessarily true *de re* by an essentialist; conversely, certain propositions commonly and uncontroversially held by the anti-essentialist to be necessarily true ***de dicto*** may be given a reading by the essentialist on which they are not necessarily true ***de dicto.***

Turning to the first result. Let 'p' name the proposition that the number of planets in the solar system is odd. The Quinean would urge that if we were *pace* Russell to construe the definite description 'the number of planets in the solar system' (henceforth 'd') as a singular term, p would not be necessarily true *de re* But the full-essentialist would reply that d serves to designate a certain object which is **essentially** the number nine, and given the definition of nine in Peano arithmetic and the axioms of Peano arithmetic, that object **must** be odd. So according to the essentialist p must be true, i.e. it is necessarily true *de re.*

Admittedly, a reading of p that makes it not necessarily true could be secured by interpreting d attributively. But the full-essentialist would reply that this would have

to be **universal** attributively. On this interpretation, p in effect says that the property of being the cardinal number of the planets is co-instantiated with the property of being odd. The full-essentialist will freely admit that there seems nothing necessary about this. Certainly the former property does not **entail** the latter. But they would urge that a **singular** attributive interpretation would still pick out an object which is **essentially** the number nine. The essentialist would concede that in contrast to a referential interpretation, a singular attributive interpretation asserts not only that the object designated is odd, but also that it is the cardinal number of the set of planets. Even so, the full-essentialist would urge that on a singular attributive interpretation, p is about the object which is **essentially** the number nine. And as the number nine, so they would insist, it must be odd, given Peano arithmetic. They would point out that the fact that this same object is only **contingently** the cardinal number of the set of planets, should not be allowed to obscure the point that what p is about is a certain object - the object is admittedly identified by way of a certain astronomical contingency - but this object is essentially a **number**, indeed essentially a specific number, and that number must be (i.e. is essentially) odd. Hence it must be true that that number is odd. Thus, the full-essentialist will conclude that p is necessarily true *de re* on a singular attributive interpretation.

Let us now turn to the second result mentioned earlier, namely that certain propositions commonly and uncontroversially held by the anti-essentialist to be necessarily true *de dicto* but not necessarily true *de re* may be given a reading on which they are not considered by the **essentialist** to be either necessarily true. *de dicto* or necessarily true *de re*. Consider for example the proposition that my younger brother is younger than I, or the proposition that my neighbour lives in my neighbourhood. Typically, either proposition will be counted as analytically true and thus necessary *de dicto*. The essentialist might add that this turns on our

reading the definite description universal- attributively. But suppose we read the definite description in both sentences as functioning referentially, as opposed to attributively. On a referential reading, the phrase 'my younger brother' serves simply to designate a certain object, and the full-essentialist would standardly urge that there seems nothing necessary in this object's being born after me (even if we endorse the Kripkean thesis that this object is necessarily related to me biologically). Similarly, on a referential reading, the phrase 'my neighbour' serves simply to designate a certain object, and there seems nothing necessary in this object's living in my neighbourhood. Thus both propositions will fail to be necessarily true *de re* on a referential interpretation of the respective definite description. But now the essentialist will continue and observe that in both sentences the import of what is stated is captured most perspicuously by a sentence of the form 'a is F', where 'a' has referential force, i.e it simply serves to denote an object. And there is nothing in the meaning of 'a is F' which analytically implies that it is necessarily true *de dicto*. So both propositions will also fail to be necessarily true *de dicto* on a referential interpretation of the respective definite description.

1.3.5 SUMMARY OF THE MODAL SEMANTICS

Let us finally gather together the results of our semantic digression, in order to determine the answer to our primary question, viz. the bearing of essentialism on systems of QML.

Firstly, where 'c' is an individual constant we have discovered that formulae of the form 'nec(Fc)' can be understood by the essentialist at least as asserting of c that it is necessarily true that it is F, while formulae of the form '(nec F)c' can be understood by them as saying that c is necessarily F, and thus that it is necessarily true that c is F. 'nec(Fc)' and '(necF)c' will be logically equivalent for non-empty

individual constants c.

Secondly, '(∃x)(nec(Fx))' and '(∃x)((necF)x)' will also be equivalent.

Thirdly, letting 'd' represent a definite description uniquely satisfied by an existent denotatum, 'nec(Fd)' and '(necF)d' will be equivalent on the non-Russellian account summarised in 1.3.3.4, the account which construes d as a singular term.. Moreover, the equivalence will hold on referential, singular attributive and universal attributive interpretations. On a Russellian analysis however, 'nec(Fd)' and '(necF)d' will not in general be equivalent, even. where the denotatum exists.

1.4 DO QUINE'S FORMAL AND INFORMAL CHARACTERISATIONS OF ESSENTIALISM COINCIDE?

We have at last reached the position where we can answer the question raised in the last paragraph of 1.2.3, viz. whether Quine's formal characterisation of Aristotelian essentialism '(∃x)((necFx & Gx) & ~necGx)' squares with his ordinary language characterisation of it. By the lights of our informal semantics, since 'x' as it occurs in the modal characterisation functions as an individual variable, its logical behaviour in quantification, taken substitutionally, is covered by the results of the previous paragraph concerning individual constants, and so the ambiguous 'necFx' will be indifferent to a reading as either '(necF)x' or 'nec(Fx)'. Similarly for 'necGx'. The following interpretation of Quine's modal characterisation emerges for the full-essentialist: some object is such that irrespective of its mode of designation it is necessarily true of it that it is F and true of it that it is G but **not necessarily** true of it that it is G. On the essentialist view, if it is necessarily true of the object in question, say x, that it is F, irrespective of mode of designation, then it could not lack the property of Fness at any time of its existence. And if it is only contingently true of x that it is G, irrespective of mode of designation, then x could lack (or could

have lacked) the property of Gness[3]. It follows on the essentialist view that Fness is an essential property of x, while Gness is non-essential, i.e. contingent. So x has some property essentially and another contingently. But this is the substance of Quine's ordinary language account of what he calls 'Aristotelian essentialism'. So the essentialist can grant that by the lights of our semantics Quine's modal characterisation of Aristotelian essentialism squares with his ordinary language characterisation. The various other forms of essentialism cited on pages 5 and 18 can now be formalised as follows:

WEAK ESSENTIALISM $(\exists F)(x)(necF)x$

STRONG ESSENTIALISM B v C

where B is $(\exists F)(\exists y)(x)(\sim Fy \ \& \ (Fx \rightarrow (necF)x))$

and C is $(\exists F)(\exists x)(Ey)((Fx \ \& \ -(necF)x) \ \& \ (Fy\&(necF)y))$

CATHOLIC ESSENTIALISM $((\exists F)(x)(necF)x \ \& \ (x)(F)((necF)x \rightarrow (y)(necF)y))$

INDIVIDUALIST ESSENTIALISM $(\exists x)(\exists y)(\exists F)(((necF)x \ \& \ -(necF)y) \ \& \ x \neq y)$

LEIBNIZIAN ESSENTIALISM $(x)(F)(Fx \rightarrow (necF)x)$

The anti-essentialist's concern of course is that an object is said to be necessarily F regardless of its mode of designation, and yet some substitutions in 'nec(Fa)' lead to a truth while others don't, even though the substituends have the same denotatum (contrast e.g. '9' and 'the number of the planets' when substituted for 'a' in 'a necessarily is greater than seven'). On a semantics other than the essentialist

3 Quine of course would refuse to accept this interpretation, in as much as it abandons a relativised *de dicto* conception of necessity in favour of an absolute *de re* conception.

semantics I have suggested, the essentialist need not adopt a referential reading of every substituend, but only of some. So e.g. '9' would be read referentially while 'the number of the planets' would not be, so that 'the number of the planets necessarily is greater than seven' would be read universal attributively and not referentially. Read in this manner the statement is of course false, just as the anti-essentialist complained it was. But then the essentialist would have to say that on such a reading the statement is not about the same object as is the statement expressed by the sentence '9 necessarily is greater than 7' on a referential reading of '9', and so the divergence in truth value between the two statements is perfectly innocuous. Such an essentialist would not talk about referential and universal attributive uses of substituends; their idiom rather is that of *de dicto* and *de re* readings of sentences. So for example a sentence with '9' as its subject term is to be read *de re*, while a sentence with 'the number of the planets' as its subject term is to be read *de dicto*. My point is however that the *de re /de dicto* contrast in this context just is that between a referential/singular attributive reading of the subject term and a universal attributive reading.

Furthermore, the fundamental diffficulties with essentialism are independent of the account of quantification appealed to. The essentialist need not adopt - as I have suggesred - a substitutional account of quantification. For ultimately what determines the truth-value of essentialist statements whether read substitutionally or objectually is a theory of essential properties. It is this metaphysical theory which ought and will be our focus of attention.

1.5 QUINE ON THE ILLEGITIMACY OF QML

My immediate concern now is to assess Quine's arguments for the contention that systems of QML are committed to Aristotelian essentialism as he explicates it, and

that since Aristotelian essentialism is philosophically suspect, QML should be proscribed. In outline, Quine's general strategy run thus: the very raison d'être of QML is to allow the possibility of quantification into modal contexts. Since such quantification is illegitimate, apparently because it requires among other things a commitment to Aristotelian essentialism, QML is equally illegitimate.

I shall first have to ascertain what Quine means by 'quantification into modal contexts' before we can determine whether: (1) such quantification is in fact illegitimate; and (2) the bearing of this on essentialism. Now one possible reason why quantification into modal contexts is illegitimate is simply that it requires a commitment to Aristotelian essentialism. Indeed that is precisely how Quine ultimately seems to want to argue. But unfortunately this line of thinking is submerged in a swamp of other considerations advanced for the illegitimacy of quantification into modal contexts, considerations which turn on the issue of the intelligibility of quantified modal locutions, considerations moreover whose dialectical positioning suggests their relevance for Quine at least to the issue of QML's commitment to essentialism, when their relevance to this issue is in fact far from clear. Thus we shall have to proceed very carefully.

1.5.1 QUANTIFICATION INTO MODAL CONTEXTS

To begin with, it is important to note that Quine believes that necessity box of modal logic is most properly understood as a monadic **predicate** rather than as an unary **sentence operator**. Specifically, the logical structure of a sentence of the form 'nec p' is most perspicuously rendered for Quine by 'p is necessary', where 'p' is a variable whose substituends are singular terms which denote propositions. Quine construes the predicate expression 'is necessary' as signifying analyticity, so that a statement of the form 'nec ... ' is true if and only if the component sentence

which 'necessarily' governs is analytic (Quine RM, pp.138-9). Of course, in *Two Dogmas of Empiricism* Quine argues that analyticity is a pseudo-concept which philosophers ought to dispense with (Quine 3GMI, p.171). Nevertheless, Quine's point is that even if we admit this concept, further devastating problems lie ahead for quantification into modal contexts

Restricting ourselves to 'necessarily' ('possibly' can be defined on the basis of 'necessarily'), and considering for the sake of simplicity monadic predicates only, we can, following Quine, understand quantification into modal contexts as either existential generalisations of the form '(∃x)necFx' or universal generalisations of the form '(x)necFx'. (Note that Quine is insensitive to the ambiguity for which we have argued of 'necF'. Needless to say, Quine would hardly set great store by this ambiguity if it were pointed out to him). Quine distinguishes such quantification into modal contexts from the following 'nec(∃x)Fx' and 'nec(x)Fx', which Quine paraphrases as ''(∃x)Fx' is analytic' and ''(x)Fx' is analytic' respectively, and are true if and only if the named sentences are analytic. Thus formulae of these forms are allowed to have, for the sake of the argument, a determinate meaning and determinate truth-conditions and indeed determinate truth-values. Any indeterminacy arises for Quine from the underlying notion of analyticity.

It is with respect to the determinacy of meaning and of truth-values that Quine contrasts 'nec(∃x)Fx' and 'nec(x)Fx' with '(∃x)necFx' and '(x)necFx' respectively, faulting the latter pair both on the count of being so indeterminate in meaning as to be unintelligible and on the count of suffering from an indeterminacy of truth-values. It is not that Quine finds such formulae ambiguous, vague or even opaque; rather he finds them incoherent. He calls them 'nonsense' (Quine RM, p.144). And even if we apply some apparently natural interpretation, Quine thinks that the truth-values

determined by such an interpretation will be problematic. But how exactly does Aristotelian essentialism figure in all of this? In the *locus classicus, Reference and Modality*, Quine writes as if all of these issues were directly germane to the question of Aristotelian essentialism. In fact it appears that in that paper Quine in effect provides us with arguments not only for the issue of a commitment to essentialism but for the unintelligibility of quantified modal locutions as well. There appear to be two such arguments. And while only one of the two arguments for unintelligibility would appear to depend on the thesis of the indeterminacy of truth-values, Quine writes as if the other argument did also. Indeed, this other argument is juxtaposed with the argument for the indeterminacy of truth-values, creating the impression that both arguments are making one and the same point, and are thus both equally relevant to the **further** issue of Aristotelian essentialism. In fact this issue is brought in at the end of RM in connection with the argument for the unintelligibility of quantified modal locutions which turns on the thesis of indeterminate truth-values. Our work as underlabourers is cut out for us then. We shall have to tease out these tangled threads in order to locate the precise place of Aristotelian essentialism in this intricate dialectic.

1.5.2 THE ARGUMENT FOR UNINTELLIGIBILITY FROM THE REFERENTIAL OPACITY OF QUOTATIONAL CONTEXTS

Of the two arguments for the unintelligibility of quantified modal locutions, let us examine the argument which is in fact independent of the thesis of indeterminacy of truth-values, but is presented as if it were integral to it. The argument is that from the referential opacity of quotational contexts. Quine introduces the important notion of a purely referential occurrence of a singular term in a sentence (Quine 3GMI, p.160). Such an occurrence is one in which the term functions simply to refer to its object (Quine 3GMI, p.160). As a criterion for purely referential

occurrence Quine invokes Frege's criterion of substitutivity of identicals, i.e. a singular term occurs purely referentially in a context if and only if substitution of any codesignative term yields a statement materially equivalent to that expressed by the original sentence. 'Failure of substitutivity reveals ... that the occurrence to be supplanted is not purely referential, that is, that the sentence depends not only on the object but on the form of the name' (Quine RM, p.133). Unfortunately, Quine says no more about the nature of this alleged dependence on the form of the name.

Quine then goes on to introduce the allied notion of a referentially opaque context: i.e. a context such that 'when, by putting a statement Ø into that context, we can cause a purely referential occurrence in Ø to be not purely referential in the whole context' (Quine 3GMI, p.161). To illustrate the notion of a purely referential occurrence let us cite Quine's example 'Giorgione was called 'Giorgione' because of his size' (Quine RM, p.136). The first occurrence of 'Giorgione' is purely referential because it passes the substitutivity test. Thus e.g. 'Barbarelli was called 'Giorgione' because of his size' is also true. But the second occurrence of 'Giorgione' is not purely referential because it fails the substitutivity test. Thus e.g. 'Giorgione was called 'Barbarelli' because of his size' is false.

To illustrate Quine's notion of referential opacity, let us examine the context '' ... '' contains just three characters'. The context is referentially opaque, for '9' occurs purely referentially in ''9 > 5', but not purely referentially in the context '' 9 > 5 '' contains just three characters', as is evidenced by the following failure to meet the substitutivity criterion: the replacement of '9' with 'the number of planets' in the truth '' '9 > 5' contains just three characters' yields the falsehood ''the number of planets > 9' contains just three characters'. (see Quine 3GMI, pp.160-161).

Quine cites the context of quotation as the referentially opaque context par excellence. This does not mean that an occurrence of a singular term within single quotes is never purely referential - for clearly the occurrence of 'Giorgioine' in ' 'Giorgione' named a chess player' is purely referential - rather that quotation can and ordinarily does destroy referential occurrence (see Quine, RM, pp.136-7).

Quine lists two other important categories of referentially opaque contexts besides quotation: those of the propositional attitudes and of modality. We have seen how Quine casts modal contexts in the mould of quotation: '9 is necessarily greater than 7' is paraphrased as ' '9>7' is analytic'. Quine finds a parallel construction for the propositional attitudes: thus 'Philip believes that Tegucigalpa is in Nicaragua' can be paraphrased as 'Philip believes 'Tegucigalpa is in Nicaragua'', but Quine thinks that while tidy, the forcing of all referentially opaque contexts into the quotational mould is unnecessary. It is sufficient for him to recognise quotation as one referentially opaque context among many (Quine RM, pp.137-8).

That modal contexts can be referentially opaque is thought by Quine to be readily evidenced by the following example: 'nec(9>5)' is true, yet 'nec(the number of planets >5)' is false, despite the truth of (the number of planets = 9' (Quine 3GMI, p.163).

Even so, why should the referential opacity of modal contexts render illegitimate quantification into them, moreover in such a way as to cast serious doubts on their meaningfulness? And what is the bearing of this on Aristotelian essentialism?

Quine argues for the illegitimacy or impropriety of quantification into modal contexts as follows: he observes that the existential quantifier (the argument could

be readily recast for universal quantifiers) in '(∃x)('six' contains 'x')' has no bearing on the character 'x' occurring within the referentially opaque context of quotation ' 'six' contains '..' ' (Quine RM, p.144). The existential quantifier is prefixed redundantly, for the work is done by the fragment ' 'six' contains 'x' '. The illusion that the character 'x' refers back to the quantifier is engendered solely by an accident of orthographic resemblance.

Unfortunately, Quine's example does not secure the point he seems to be wanting to make, viz. that quantifiers outside a referentially opaque construction need have no bearing on what is **intended** to function as a variable inside it (Quine RM, p.144), that the **intended** variable cannot be reached by the quantifier so to speak (Quine RM, p.150). For the character 'x' in 'six' clearly does not function as a variable, and the second occurrence of 'x' does not either. But it is not even **intended** to: this character is meant to function as the twenty fourth letter of the alphabet. Since these two occurrences of 'x' are the only even remotely plausible candidates for variablehood, the existential quantifier governs no variables. It follows, albeit vacuously, that the quantifier has no bearing (does not reach) any variables inside, for there are none.

Elsewhere (Quine RM, p.141) Quine does provide the kind of example appropriate to securing his point, namely that quantifiers outside a referentially opaque construction need have no bearing on **intended** variables inside it. Consider the context ' ' ... ' contains six letters': it is referentially opaque. Now an existential generalisation on ''Cicero' contains six letters' could yield two different sentences, viz.

(1) (∃x)(x contains six letters)

or

(2) (∃x)('x' contains six letters)

(1) would be an existential generalisation of the context ' ... contains six letters', whereas (2) would be an existential generalisation of the context '' ... ' contains six letters', which is the context that concerns us.

Quine recasts '(∃x)('x' contains six letters)' as 'there is something such that 'it' contains six letters' or as ' 'something' contains six letters' (Quine RM, p.141). On either casting, the sentence expresses a falsehood, for there is nothing such that 'it' contains six letters, since it is simply false that the word 'it' contains six letters, and similarly it is false that 'something' contains six letters. The quantifier '(∃x)' turns out to be redundant, and indeed for the same reason that the quantifier proved to be redundant in the original example of '(∃x)('six' contains 'x')', namely that 'x' in ' 'x' ' does not function as a variable. Of course in the original example 'x' was not intended to function as a variable, but in the present case the intention just is for it to function as a variable. In either case, the context bound by the quantifier lacks variables altogether. So by default, it follows in either case that the quantifier outside does not 'reach' any variable inside, for there are none. The only difference between the two cases is that the original example consists of a truth preceded by an irrelevant quantifier, while the present example consists of a falsehood preceded by an irrelevant quantifier. Now the context ' ' ... ' contains six letters' exhibits the logical structure ' F ' ... '.' And in general, contexts exhibiting such a logical structure are such that by quantifying on a singular term occurring within the quotation we obtain a variable-free context preceded by an irrelevant quantifier. But according to Quine modal contexts like 'necGa' are most perspicuously paraphrased as ' 'Ga' is analytic', where ' 'Ga' ' is a singular term designating the proposition that Ga. And ' 'Ga' is analytic' exhibits the logical structure ' F ' ...' '(Consider 'F' as 'is analytic'). Therefore quantifying into ' 'Ga' is analytic' will produce only an irrelevant quantifier preceding a variable-free context. Since ' 'Ga'

is analytic' is equivalent to 'necGa', to quantify into 'necGa' in the manner of '(∃x)necGx' is in effect to indulge in the appending of irrelevant quantifiers that have no variables to bind; it is a degenerate case of quantification.

Such then appears to be Quine's argument that the referential opacity of modal contexts renders quantification into them illegitimate. But if this is the argument, it seems overstated to say it has established its conclusion. It is one thing to say that the quantification in question is degenerate because redundant. Admittedly it is in this sense that one can endorse Quine's claim that quantification into modal contexts is improper (Quine RM, p.143) or unwarranted (Quine RM, p.143) or loses its justification (Quine RM, p.140). But is redundancy the same as logical impropriety in the strong sense of **impropriety** Quine appears to intend? Certainly **standard** formation rules permit redundant non-modal quantifications (consider e.g. '(x)Fa').

More importantly, how does the redundancy or pointlessness of quantification undermine its intelligibility? Quine's contention after all is that such quantification is at worst nonsensical or at best productive of unintended sense (Quine, RM, p.143). Quine concedes that '(∃x)('x' contains six letters)' consists of a **falsehood**, viz. ' 'x' contains six letters' preceded by an irrelevant or redundant quantifier, viz. '(∃x)', and so is itself a falsehood. And a falsehood, though false, is **meaningful**. Quine concedes as much when he calls this falsehood 'an unintended **sense**' of quantification into referentially opaque contexts. But he will then have to concede that existentially quantifying on 'a' in 'necGa', which is ' 'Ga' is analytic' by Quine's lights, will produce the **falsehood** '(∃x)('Gx' is analytic), consisting of the **meaningful** but false ''Gx' is analytic' preceded by an irrelevant quantifier. The claim that quantification into modal contexts produces **nonsense** is far from established by the argument for the redundancy of quantifications into referentially

opaque contexts. Admittedly, nothing is left of QML if all its quantifiers are redundant, and if what is expressed by the quantification is unintended and false, but this is a separate point from that of nonsensicality.

It might be thought that we are conceding too much to Quine in allowing his paraphrase of 'necFa' as ' 'Fa' is analytic'. And if we disallow this move, we might then think ourselves to have blocked the conclusion that quantification into modal contexts is idle or degenerate. Unfortunately, the problem is not eliminated if we do disallow the paraphrase in terms of 'is analytic'. For an alternative paraphrase is always available in terms of the neutral 'is necessary'. Surely no one would contest the equivalence between 'necFa' and ' 'Fa' is necessary' or ' 'Fa' necessarily is true'. But the context ' ' ... ' is necessary' equally leads to degenerate quantification. I think that a proponent of the **non-degeneracy** of quantification into modal contexts will have to take the tack of arguing that while 'necFa' and ' 'Fa' is necessary' express equivalent propositions, the **logical** structure of the proposition expressed will be better exhibited by 'necFa' than by ' 'Fa' is necessary', and that we have to hand an informal account of what '(∃x)necFx' means, namely that something is such that it necessarily is F; and this is not a degenerate quantification. This then would explain how essentialism would have to enter the picture.

1.5.3 THE ARGUMENT FOR UNINTELLIGIBILITY FROM THE INDETERMINACY OF TRUTH-VALUES OF QUANTIFIED MODAL CONTEXTS

Immediately following on in *Reference and Modality* from the argument from the referential opacity of quotational contexts is an argument for the indeterminacy of the truth-values of quantified modal contexts. Yet Quine writes as if the observations he has made in connection with the former argument bear directly on

the latter (see Quine RM, infra p.141). In fact the issue of referential opacity does connect up with the issue of indeterminate truth-values, but as we shall see this has nothing to do with the referential opacity of **quotational** contexts. The primary aim however is the same: Quine wants to argue that quantifying into modal contexts presents serious problems of interpretation (Quine RM, p.142). This time he appeals to the indeterminacy of truth-values to establish his contention.

Unfortunately, the thrust of Quine's argument is not altogether clear, even if we detach it from its confusing association with the previous argument from quotational contexts. Our reconstruction of Quine's argument, for what it is worth, is as follows.

Take a sentence such as 'nine is necessarily greater than seven'. If we quantify (by existential generalisation say) on the occurrence of 'nine', we get '($\exists$x)(x is necessarily greater than 7)'. This is true according to the standard objectual interpretation of quantificational theory (Quine of course dismisses the substitutional interpretation in as much as it is not the final court of appeal in determining the ontological commitments of a theory), just where some object satisfies the predicate or corresponding condition 'is necessarily greater than nine'. Given this perfectly straightforward truth-condition, the sentence in question will be either true or false but not both in a classical two-valued logic, the logic within which not only Quine but the majority of mainstream philosophers operate. By attaching an existential quantifier to an open sentence corresponding to a particular predicate, we are saying that that predicate is satisfied by something, and thus that some thing is so and so or has a certain property (Quine of course would balk at such talk of properties, insofar as it carries realist connotations). So proper (or legitimate) existential quantification requires that there be some **object** that of itself

and in itself and irrespective of its means of designation or specification has a certain property - Quine calls this object 'the neutral thing' (Quine RPM, p.183). And given a bivalent logic endorsing the law of excluded middle, a proper or legitimate existential generalisation will be determinately and unequivocally true or false. So that if a putatively proper existential generalisation fails to meet the requirement of being unequivocally true or false, then the generalisation in question cannot be proper after all. The quantification would fail as a genuine indefinite predication, i.e. a predication of a genuine predicate (condition/property) of some (neutral) object, for genuine predication of an object is insensitive to the mode of that object's specification. So the predicate will fail to express a genuine property and thus fail to characterise any object; alternatively, the predicate will express a genuine property but will fail to be predicated of any particular object. On either alternative, the existential generalisation will fail to make an assertion, either because it says nothing about something, or because it says something about nothing; in either case, it fails to say anything and is thus meaningless, resisting coherent interpretation as Quine would say (Quine RM, p.149). Since universal generalisations are definable in terms of existential generalisations, it follows that quantification *tout court* into a context, such that the result fails of semantic univocity, will be **meaningless**.

Now by a very simple argument Quine urges that '(∃x)(x necessarily is greater than 7)' fails to be unequivocally true or false: the object which satisfies the putative condition 'is necessarily greater than 7' on designation as the square of 3 say, happens to violate it on designation as the number of planets in our solar system, since 'the number of planets in our solar system necessarily is greater than 9' would appear to be false (Quine RM, p.142). It is precisely this alleged failure of '(∃x)(x necessarily is greater than 7)' to be insensitive to the mode of designation of the neutral object x that renders the sentence meaningless. It follows that necessarily

being greater than seven is not a trait of a neutral object, but depends rather on the manner of referring to that object (Quine RM, p. p.142). Quine concludes that necessary greaterness than seven (or necessary fulfilment of ' is greater than seven') makes no sense as applied to an object x, for the necessity attaches to instances of 'x > 7' only relative to the particular means of specifying x, say as '3', as opposed to specifying it as 'the number of the planets' (Quine RM, p.143). Generalising from this example, Quine urges that being necessarily so and so (and derivatively, **possibly** so and so), is not a trait (property) of the object concerned, but depends rather on the manner of referring to it (Quine, RM, p.142). And so it is nonsense to say neutrally that there is **something,** x, that necessarily is so and so (Quine RPM, p.184).

1.5.3.1 RESPONSE TO THE ARGUMENT FROM THE INDETERMINACY OF TRUTH-VALUES

In response to Quine's reconstructed argument we make four general points.

Firstly, the ruling out of court of putative predicates such as 'necessarily is greater than seven', is part of Quine's general strategy to preserve the extensionality of first order logic. Indeed, this concern seems a primary motivation for Quine's general attack on QML.

Secondly, the notion of referential opacity also plays a role in this argument. In fact, with the aid of this notion we achieve an explanation, amenable to Quine's framework, of the impropriety of quantification into modal contexts. Since the occurrence of the singular term 'a' in the context 'necFa ...' is not purely referential, 'a' fails to function **simply** to designate a (neutral) object. What instead or in addition its functions are is never elaborated on by Quine - an unsatisfactory lacuna

indeed. But perhaps we can fill in the details in a vein faithful to what Quine probably had in mind, as follows: since the logical function of 'a' in 'necFa' is not purely designative, a casting of the logical structure of 'necFa' within a simple subject - predicate mould such as '(necF)a' where 'necF' is the predicate and 'a' the subject, would fail to capture the more complex logical function of 'a', and thus would perhaps seriously misrepresent the logical structure of 'necFa'. *A fortiori* the existential generalisation of 'necFa' would also go astray. This is not to say that the correct logical structure of 'necFa' could not be of subject-predicate form, merely that if so, it would probably be more complex; perhaps the necessity operator would be contextually eliminated in the proper reconstruction. To that extent 'nec ... > 7' would not express a genuine trait or property of an individual, and so at worst the expression would be ill-formed and improperly predicable of an object, in other words a spurious predicate, and at best it would constitute a definitional abbreviation for some more complex expression, thus being contextually eliminable (cf. on this point Russell's treatment of existentials with names and definite descriptions - albeit in Russell's treatment, the reconstruction does not retain a subject-predicate form). That my account probably reflects Quine's own thinking is corroborated by the following extract:

> we can reasonably infer '($\exists$x) nec x>5)'from 'nec (9>5)' only if we regard the latter as telling us something about the *object* 9, a number, viz. that it necessarily exceeds 5. If 'nec(... >5)' can turn out true or false 'of' the number 9 depending merely on how that number is referred to ... then evidently 'nec(x>5)' expresses no genuine condition on objects of any kind. If the occurrence of '9' in 'nec(9>5)' is not purely referential, then putting 'x' for '9' in 'nec(9>5)' makes no more sense than putting 'x' for 'nine' within the context 'canine'.' (Quine 3GMI, pp. 172-173.)

My third general point concerning Quine's reconstructed argument is that if we grant the essentialist a *de re* conception of necessity, then the argument breaks down at the crucial claim that '(Ex)(nec x>7)' fails to admit unequivocally of a single

truth-value. To vindicate this point we will appeal to the semantic apparatus explained earlier, and invoke the distinction between singular (referential or attributive) uses of definite descriptions and universal attributive uses. To recall: Quine's argument turns on the observation that 'nec 9>7' is true, while 'nec the number of the planets > 7' is false. So it is indeterminate whether **something** is indeed necessarily greater than seven. Now on a **referential** interpretation of 'the number of the planets' (henceforth 'n'), 'n' simply designates that neutral object that also happens to be the number nine, but then for the **full-essentialist** 'nec n>7' will be true, not false, *pace* Quine, for if it **is** essential to this neutral object's identity that it be both a number (as opposed to a non-number), and furthermore that it be a certain number, to wit the number nine (and not some other number), then given Peano's axioms it would surely have to be true (it would be necessarily true) of this object that it is greater than seven, in which case it follows that this object necessarily is greater than seven. Far from '(∃x)(nec x>7)' 's failing to admit unequivocally of a single truth-value, it is straightforwardly true for both the substitution instance 'nec 9>7' and 'nec n>7', on a **referential** interpretation of 'n'. The same can be easily seen to hold for a **singular attributive** interpretation of 'n'. The anti-essentialist Quine of course will not wear the use of the phrase 'essential to the neutral object's identity'.

What of a **universal attributive** interpretation of 'n'? This is equivalent to '(x)(x = n → nec x>7)', which clearly need not be true; in fact, it is false, for surely there is no necessary connection obtaining between the property of numbering the planets and that of being greater than seven, in which case there will be falsifying instances of '(x)(x = n → nec x>7)', i.e. objects which could number the planets but would not be greater than seven, and hence certainly not necessarily greater than seven. Is Quine vindicated then? No, for his claim is that '(∃x)nec x>7' is true for neutral x when it

is identical with (Peano) nine but false for the **same** neutral x when it is identical with the number of the planets, where the value of 'x' is one and the **same** neutral object. But clearly '9' does not designate the one and the **same** object as is designated by an **attributive** use of 'n': 'n' designates, if anything, a property or class. Even if it is retorted that nine, being a number, is itself properly speaking a property or class, surely it is not the **same** property or class as that which 'n' would have to designate on the universal attributive view. Thus on all possible interpretations of 'n' as a singular term, 'n necessarily-is-greater-than-seven' is straightforwardly true (I have omitted consideration of Russellian interpretations of this sentence, for they do not treat 'n' as a singular term)[4].

As a postscript on my third general point concerning Quine's reconstructed second argument for the impropriety or illegitimacy of quantification into modal contexts, I observe that my discussion of the third point focussed on the issue of '(∃x)(nec x>7)"s having **one determinate truth-value**. As a corollary of my investigation substitutivity *salva veritate* is seen to hold. But this **corollary** was not the focus of my discussion of the third point. Interestingly, many philosophers free of Quine's misgiving about quantification into modal contexts, **have** focussed on the conclusion of this corollary, which in effect is tantamount to the issue of the extensionality of QML. Ravnkilde is typical of such philosophers; he understands Quine to be arguing as follows:

(A) Existential generalization is permissible only from statements embodying contexts that express properties;

(B) A singular context expresses a property only if its blank has a referential position

---

4 See Appendix A of this chapter.

(C) The blank in such a context has referential position only if substitutivity holds in that context

(D) Substitutivity does not hold in the context 'necessarily, ... is greater than 7' (see Ravnkilde, pp.53-54).

Ravnkilde's (A) reminds us of the second of Quine's arguments, viz. the argument under discussion in general points (1) to (3) on pages 43-45. But Ravnkilde's (B) signals a subtle swing away from the direction of this reconstructed argument. For our reconstructed argument turns on the failure of a predicate context to express a property because of the **indeterminacy in truth-value** of the resulting existential generalisation, whereas Ravnkilde's (B) turns on the failure of a predicate context to express a property because of the **non-pure-referentiality of singular terms** occurring in such a context. Given his endorsement of Quine's criterion of referential occurrence in terms of subtitutivity *salva veritate*, we can see readily why Ravnkilde perceives a need to rebut (D) in order to salvage QML. But it is my contention that Ravnkilde has reconstructed a straw-man in the argument (A) (B) (C) (D). Admittedly, the textual evidence of Quine's *Reference and Modality* would appear to confirm Ravnkilde's interpretation, in as much as Quine juxtaposes an argument from the referential opacity of quotational contexts (thus (C)→(D)) with that from the illegitimacy of quantification into modal contexts because of indeterminacy in truth-value of the resulting generalisation (thus (A)). This textual juxtaposition of arguments suggests a dialectical unity of argumentation; Ravnkilde then goes on to secure this unity through a mediating or bridging principle (B)[5]. I

---

5 (B) is plausible, irrespective of whether it properly belongs to the argument or not. For the notions of singular term (subject) and singular context (predicate) are correlative, so that if the (singular term) blank fails to be purely referential, i.e. fails just to refer to the subject,

have however pointed out two **separate** arguments in Quine for the illegitimacy or impropriety of quantification into modal contexts, one from the referential opacity of quotational contexts, the other from the indeterminacy in truth-value of the resulting generalisation. Ravnkilde's reconstruction conflates these separate arguments, integrating them through premiss (B). In effect Ravnkilde appropriates the major thesis of the argument from indeterminacy, viz. that (existential) generalisation into modal predicate contexts is illegitimate because of the failure of such contexts to represent genuine properties, and conjoins it with the major thesis of the argument from the opacity of quotational contexts, viz. that modal contexts are referentially opaque because of the failure of substitutivity *salva veritate*. This conjunction of theses generates the impression that failure of substitutivity bears directly on the failure of modal predicate contexts to express genuine properties. Consequently, Ravnkilde is led to focus on the issue of **substitutivity** whereas we focussed on the issue of indeterminacy in truth-value. From the point of view of **rebutting** Quine I could have proceeded with Ravnkilde to interpret Quine as arguing that a predicate fails to express a genuine property in as much as it is a **referentially opaque** context[6], but equally I could have proceeded, as I in fact did, to interpret Quine as arguing that a predicate fails to express a genuine property because (existential) generalisation into the predicate context suffers from an indeterminacy of truth-value. My procedure appears more faithful to the thrust of the text, notwithstanding Quine's own confusion on the point.

With this postscript on my third general point behind me, I now finally come to make my fourth and final general point concerning my reconstruction of Quine's

---

then the singular context enclosing the blank correlatively fails to express a property that is predicated of the would-be subject.

6 See appendix A of this chapter.

second argument for the impropriety of quantification into modal contexts. That is the question of just how exactly the argument is supposed to bear on the issue of a commitment of QML to essentialism. To quote Quine at length:

> the only hope of sustaining quantified modal logic lies ... in arguing or deciding that quantification into modal contexts makes sense even though any value of the variable of such a quantification be determinable by conditions that are not analytically equivalent to each other. The only hope lies in accepting ... [both (32) $x = 3\sqrt{x} \neq \sqrt{x}$ and (33) there are exactly x planets] and insisting, despite it, that the object x in question is necessarily greater than 7. This means adopting an invidious attitude toward certain ways of uniquely specifying x, for example (33), and favoring other ways, for example (32), as somehow better revealing the 'essence' of the object. [Satisfying (32) and c]onsequences of (32) can, from such a point of view, be looked upon as necessarily true of the object which is 9 (and is the number of the planets), while [satisfying (33) and] some consequences of (33) are rated still as only contingently true of that object ... this reversion to Aristotelian essentialism ... is required if quantification into modal contexts is to be insisted on. An object, of itself and by whatever name or none, must be seen as having some of its traits necessarily and others contingently, despite the fact that the latter traits follow just as analytically from some ways of specifying the object as the former traits do from other ways of specifying it.
>
> (Quine RM, pp.147-148)

The gravamen of Quine's second argument against QML is that modal quantifications suffer from an indeterminacy of truth-value. This in turn is because Quine construes necessity not as *de re* but as *de dicto* or metalinguistic: the **statement** that nine is greater than seven is necessarily true, the **statement** that the number of the planets is greater than seven is not necessarily true, even though both statements are about the same subject. Now Quine concedes that the full-essentialist can meet his objection head on by renouncing a *de dicto* interpretation of necessity and adopting a *de re* interpretation in its place. Since both statements are about the same subject for the full-essentialist, they must both have the same truth-value, in this case truth. So contrary to Quine modal quantifications do not suffer from an indeterminacy of truth value: If the property predicated in a statement is essential to the subject of that statement, then any statement predicating that

property of the subject will be necessarily true, regardless of the manner in which the subject is linguistically identified. Similarly, if the property in question is not essential to the subject, then any statement predicating that property or the subject will not be necessarily true, regardless of the manner in which the subject is linguistically identified.

> This is how essentialism comes in: the invidious distinction between some traits of an object as essential to **it** (by whatever name) and other traits of it as accidental. I do not say that such essentialism, however uncongenial to me, should be uncongenial to the champion of quantified modal logic. On the contrary, it should be every bit as congenial as quantified modal logic itself (Quine RPM, p.184).

That essentialism is uncongenial to Quine is made most clear when he calls it 'a metaphysical jungle' (Quine 3GMI, p.176) and 'an unreasonable philosophy' (Quine RM, p.149).

We may now finally observe how Quine's second argument for the impropriety of quantification into modal contexts bears on the issue of QML's commitment to essentialism: the proponent of QML can evade the charge of truth-value indeterminacy only on pain of endorsing the allegedly incoherent, viz. essentialism.

1.5.4 IS QML COMMITTED TO ESSENTIALISM?

Are systems of QML committed to essentialism then? The answer is simply that they are. Semantical considerations show why: locutions of the form 'Fa' (Quine wants of course to disallow singular terms) or '($\exists$x)Fx' or '(x)Fx' are standardly interpreted on the **objectual** interpretation (which Quine of course endorses) with respect to a domain of **neutral** objects, i.e. objects in themselves in abstraction from their various (if any) modes of designation or specification. Thus for example '($\exists$x)Fx' will be true just where some neutral object however (if at all) it is described or named, is F. Similarly, '($\exists$x)necFx' or '(x)necFx' will be true on an interpretation

just where some or every neutral object x in the domain necessarily-is F, i.e. some property F is essential to x. This then is the proper force of Quine's claim that systems of QML are committed to Aristotelian essentialism. In general objectual quantificatory semantics presuppose the meaningfulness of a distinction between the essential (necessary) and accidental (contingent) properties of an object qua that object, i.e. absolutely and irrespective of mode of designation or presentation. Whether an object does in reality have a certain property essentially is a separate question which if answered affirmatively finds expression in some system of QML or other. The significant point is that common to these systems is the thesis that a distinction between **absolutely** essential and accidental properties is intelligible. QML just is after all that logic which enables such a distinction to be expressed and thus **is** indeed committed to essentialism (or what Quine calls 'Aristotelian essentialism')[7].

---

7 In the light of the fact that Gupta has devised the systems of QML LI→L4, which are neutral on the question of essentialism, being committed neither to its meaningfulness nor to its meaninglessness, it is not entirely correct to say that the champion of QML **must** settle for essentialism (see Gupta, pp. 90-91).

As for counterpart-theoretic treatments of QML, strictly speaking there is no need for the underlying semantics to be construed as appealing to the notion of an object's having necessary (or accidental) properties, in as much as an object in counterpart-theoretic semantics is **world-bound**, i.e. it does not appear in other worlds. Nonetheless, the whole point and spirit of counterpart-theory just is to make possible the expression of *de re* necessity and possibility.

1.5.5 QUINE'S OBJECTIONS TO ESSENTIALISM

1.5.5.1 ESSENTIALISM CONFLICTS WITH THE EXPLANATION OF NECESSITY THROUGH ANALYTICITY

For Quine of course, the notion that an object can necessarily-be F *qua* G is in order, in so far as a relation of analytic entailment holds between being G and being F. But to talk of an object's necessarily-being F qua itself, i.e. absolutely, is for Quine senseless. Thus necessarily exceeding seven is true of the number nine relative to a specification of it as 'the successor of eight', but not relative to a specification as 'the number of the planets.' In abstraction from any specification or description, it is simply meaningless to Quine to talk of the neutral thing itself - which is a number, the number nine and the number of the planets - as necessarily exceeding seven. It is nonsense to say neutrally that there is **something,** x, that necessarily exceeds 7 (Quine RPM, p.184). Even if, contrary to fact, numbering the planets analytically entailed exceeding seven, the point for Quine would still be that the number of planets necessarily exceeds seven only relative to its specification as the number of the planets, and not independently of any specification. Similarly, the number nine necessarily exceeds seven only **relative** to a specification such as 'x = $3\sqrt{x} \neq \sqrt{x}$', and not **absolutely** - this point, Quine would probably urge, tends to be overlooked, in as much as we tend to think of nine as **absolutely** necessarily exceeding seven, perhaps because the (Peano) mathematical specification of nine, predominant as it is, tends to supplant the non-mathematical specifications and become identified with the neutral object that is the number nine. For Quine, it would just be meaningless to speak of this neutral object's essentially being a **number** qua itself, let alone its essentially being a number that exceeds seven qua itself.

To recapitulate: for Quine, x essentially is F only *qua* some G, and x is essentially F *qua* G only if being G analytically entails being F. QML of course opposes this line of thinking in three respects:

(1) it allows for the expression of the possibility that a neutral object a may have a certain essential property absolutely, i.e. a essentially is F *qua* a;

(2) furthermore it allows the expression of an inegalitarian stance toward the various analytically independent ways of specifying the object in question: thus one

> of the determining traits, the successor of 8, is counted as a necessary trait of the number [nine]. So are any traits that follow from that one, notably the exceeding of 7. Other uniquely individuating traits of the number [nine], notably its numbering of planets, are discounted as contingent traits of the number and held not to belie the fact that the number does still necessarily exceed 7 [*qua* itself.] (Quine RPM, p.184.)

(3) finally QML does not require that any relation of analytic entailment obtain between the circumstance that a is what it is, viz. a, and the circumstance that it essentially is F.

Quine goes on to urge that essentialism conflicts with the empiricist program of explaining necessity through analyticity.

> For the appeal to analyticity can pretend to distinguish essential and accidental traits of an object only relative to how the object is specified, not absolutely. Yet the champion of quantified modal logic must settle for[non-relative, i.e absolute] essentialism. (Quine RM, p.148.)

As an *ad hominem* against advocates of QML, this argument presupposes that these advocates will endorse the empiricist program of explaining necessity through analyticity. Setting this point aside, why exactly must an appeal to analyticity pretend to determine essential properties only relative to a specification? As Quine epigrammatically declares: 'necessity resides in the way in which we say things, and

not in the things we talk about' (Quine 3GMI, p.176). But this would hardly **explain** why an appeal to analyticity must determine essential properties only relative to a specification; it would be at best a counterclaim against the essentialist who would urge that essential properties **can** be enunciated without relativisation to a specification.

Nonetheless it is not difficult to see how necessity understood as analyticity must determine essential properties only relative to a specification. For if 'necp' is true only if 'p' is analytically true, then 'nec(Fa)' is true only if 'Fa' is analytically true. Now it might be that 'Fa' is analytically true while 'Fb' is not, even though 'b' refers to the same individual as does 'a', in which case F is necessary to the individual in question only in so far as it is designated as 'a', but not when designated as 'b'. Thus on a view of necessity as analyticity necessary greaterness than 7 makes no sense as applied to a **number** x; necessity attaches only to the connection between 'x > 7' and the particular method ... [x = 3$\sqrt{x}\neq\sqrt{x}$] as opposed to ... [there are exactly x planets] of specifying x (Quine RM, p.143).

Suppose the full-essentialist were to endorse the empiricist equation of necessity with analyticity. On this view a neutral object designated by a neutral singular term 'a' would be essentially F just where 'a is F' is analytically true. What predicates 'F' would satisfy this condition? Clearly, 'is identical with a'; 'is either-red-or-not-red' ; 'is divisible by two if even' etc. But these are bloodless counterfeits of the **bolder** examples standardly advanced by the full-essentialist, viz. having a certain origin or falling under a certain sortal or belonging to a certain natural kind. In fact, to make good these bolder examples the full-essentialist will freely have to renounce the empiricist equation of the necessary with the analytic, thereby opening the door to synthetic necessary truths. Whether proper names in fact have a sense

allowing some sentences of the form 'Fa' to count as analytically true is an issue that will occupy us in chapter .

More importantly, 'a is identical with a' and 'a is either-red-or-not-red' appear to be analytically true and thus to have a determinate truth-value, viz. truth. But then what remains to bar a concession from Quine that here we have **absolutely** essential properties *tout court*, and not merely essential properties relative to a specification? For 'a' here functions merely to denote the neutral object independently of any particular specification of that object. I think Quine would have to reply that 'is identical with a' and 'is either-red-or-not-red' are **pseudo**-predicates. But it is not at all clear how he would distinguish pseudo-predicates from genuine predicates, certainly not on the basis of pseudo as opposed to genuine **properties.** For Quine renounces properties as he does all intensional entities, i.e. entities whose individuative conditions depend on meaning.

In any case, the full-essentialist who subscribes to a bolder kind of full-essentialism will have to deny Quine's crucial assumption that necessity is *de dicto*. For even if necessity is not tantamount to analyticity, a *de dicto* conception of necessity would still conflict with absolute essentialism, in so far as a 'bolder' predication of an object might be necessarily (but not analytically) true when the object is referrred to in one way, while the same predication of the same object might fail to be necessarily true when it is referred to in another.

Since it is the very **meaningfulness** of the notion that a neutral object has a certain property essentially (or - derivatively - contingently) qua that neutral object, which Quine is suspicious of, Parsons misses the point - as far as Quine is concerned - when he declares that the commitment of QML to **some** forms of essentialism, viz

weak essentialism as opposed to strong essentialism, is acceptable because it is innocuous. Parsons here understands 'commitment of QML to essentialism' not so much in the sense of 'allowing the expression of formulae like '(∃x)necFx' but rather in a proof-theoretic sense: to say a system of QML is committed to essentialism in Parsons' sense is to say that certain kinds of modal formulae E [8] are derivable either directly from the axioms of the modal system in question or from their conjunction with certain uncontroversial non-modal facts (see Parsons, EQML, p.78). On this understanding of a commitment to essentialism, Parsons argues that most (all?) systems of QML are committed to a weak form of essentialism, for formulae of the form

'$(\exists x_1) \ldots (\exists x_n)nec(Ax_1 \ldots x_n V \sim Ax_1 \ldots x_n)$' or '$(\exists x)necIxx$',

where 'I' is the identity predicate, are provable in most axiomatisations of QML. Such formulae correspond of course to a weak essentialism, in as much as **every** object xi is such that it has the essential property of being either A-related to $<x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_n>$ or not, or of having essentially the property of self-identity.

Parsons contrasts commitment to weak essentialism with commitment to strong essentialism, in particular that form of strong essentialism the theorems of which state that a property is shared by two different objects but is essential to one of them only. Properties expressed by troublesome predicates such as 'is rational', 'has ten protons', and Quine's 'is greater than seven' are what Parsons has in mind here. Using model-theoretic results, Parsons concludes that no standard system of QML is committed to strong essentialism, in as much as sentences of the form '(∃x)nec x is rational' or '(∃x)nec x>7' are not derivable either from the axioms alone or from

---

8 The exact specification of 'E' varies with the kind of essentialism to which the commitment is alleged)

their conjunction with certain uncontentious non-modal facts (see Parsons, EQML, pp.78-82)

Parsons finds a commitment (in his sense) to weak essentialism acceptable because it is innocuous, whereas he finds strong essentialism troublesome. Happily, in Parsons' eyes, QML is not committed to strong essentialism. Parsons finds untroublesome the possibility of a particular property's being instantiated necessarily provided that every one of its instances instantiates it necessarily. By corollary, Parsons finds acceptable the notion that a certain property be instantiated contingently by a certain individual, as long as **every** individual instance of the property instantiates it only contingently. Thus Parsons accepts as unproblematic that sentences like '(∃x)nec(Ax v~Ax)' or '(∃x)necIxx' count as theorems of most systems of QML.

What Parsons does find problematic is the possibility of the one property's being necessarily instantiated by one individual and only contingently by another. Thus Parsons rejects (as meaningless) sentences like '(∃x)nec x>7' or '(∃x)nec x is rational'. Parsons thinks that the only satisfactory account of a property's being instantiated necessarily must reside in the nature of the property itself, or of its corresponding predicate, rather than in the nature of the individual instance (Parsons, p.76). It is precisely the denial of this point which seems to underlie the more interesting or bold forms of full-essentialism, those forms namely that hold some properties to be essential only to some of their instances. Such forms of full-essentialism appear to resist an account of necessary instantiation in terms of the properties themselves. Thus there would have to 'be something about the object which gives rise to the necessity. But what could this be? The lack ... of a

satisfactory answer ... makes this version of essentialism a real source of philosophical perplexity' (Parsons EQML, p.76).

But Parsons has missed Quine's point. It is immaterial to Quine whether the one property is universally essential or essential to only some of its instances. The important point is that is is simply incoherent to talk of that property's being essential to a particular instance **absolutely**, i.e. independently of its mode of specification or designation. Thus Quine would find weak and strong forms of essentialism equally problematic.

1.5.5.2. ESSENTIALISM LEADS TO INCONSISTENCY

In another attempt to discredit essentialism Quine tries primarily to induce a sense of irresolvable puzzlement allegedly attendant on the adoption of the essence/accident dichotomy:

> Perhaps I can evoke the appropriate sense of bewilderment as follows. Mathematicians may conceivably be said to be necessarily rational and not necessarily two-legged; and cyclists necessarily two-legged and not necessarily rational. But what of an individual who counts among his eccentricities both mathematics and cycling? Is this concrete individual necessarily rational and contingently two-legged or vice versa? Just insofar as we are talking referentially of the object, with no special bias towards a background grouping of mathematicians as against cyclists or vice versa, there is no semblance of sense in rating some of his attributes as necessary and others as contingent. Some of his attributes count as important and others as unimportant, yes, some as enduring and others as fleeting; but none as necessary or contingent'. (Quine WO, p.199).

How is Quine's argument supposed to work exactly? Supposedly, he wants to highlight a puzzling inconsistency resuling from the essence/accident dichotomy. If the inconsistency is real, then there must be something amiss with the dichotomy. In fact, the inconsistency is supposedly two-fold: a certain concrete individual x, being

a mathematician, necessarily is rational, but also not necessarily rational, since x is also a cyclist; similarly x, being a cyclist, necessarily is two-legged, but also not necessarily two-legged, since x is also a mathematician. Let us address the former inconsistency and in particular an approach by the essentialist to dissolve it. This approach will apply *mutatis mutandis* to the latter inconsistency. The argument runs thus:

(1) mathematicians necessarily are rational.

(2) cyclists are not necessarily rational

(3) x is a mathematician and a cyclist

Therefore

(4) x necessarily is rational and it is not the case that x necessarily is rational.

Now an essentialist, i.e. *de re* reading of (1) would be : 'every mathematician is such that that mathematician necessarily is rational'. The gist of (2) on an essentialist reading would appear to be: ' every cyclist is such that it is not the case that that cyclist necessarily is rational '. Given these readings of (1) and (2), (4) is seen to follow from (1) (2) (3). But then it can be challenged whether the full-essentialist would concede the truth of either (1) or (2) on an essentialist reading. Waiving (2), let us consider a possible counter-instance for the full-essentialist to (1). Was, for example, that famous concrete individual, the mathematician Gauss, **necessarily** rational? *Qua* mathematician he could hardly be conceived not to have been rational, but could he qua that concrete individual be conceived not to have been rational? The answer according to the pre-analytic intuitions to which the full-essentialist appeals is surely that he could be; could he not after all have been born an imbecile? If so, then (1) admits of a counterinstance and so is false; the argument for the inconsistency then fails. (The essentialist's parallel argument against Quine's argument for the second of the two inconsistencies would question whether a certain individual who counts cycling among their hobbies could have been born

with one leg only; the answer according to our pre-analytic intuitions is surely affirmative.) To admit that Gauss could have been born an imbecile is to deny that he necessarily was a mathematician. But this does not conflict with (3). If (3) had read 'x necessarily is a mathematician and is a cyclist' then Gauss would have had to be rational, if only because being a mathematical genius requires being rational. In that case, there would indeed have followed an inconsistency, in fact two-fold: Gauss both would and would not have been necessarily a mathematician, and he both would and would not have been necessarily rational. But as things stand, the essentialist need not hold that Gauss **had** to be a mathematician-they would urge that he could surely after all have been an illiterate cobbler. I conclude that the inconsistency to which Quine adverts dissolves if the essentialist is allowed to read (1) and (2) as I suggested they would want to (cf. Plantinga 1974, pp.24-26, for similar observations)[9]. That is not to say that the essence/accident dichotomy is thereby divested of all puzzlement. Why, after all, is it that Gauss might not have been a mathematician? And why is it in contrast that Gauss had to be human, as the full-essentialist will often urge? These are questions that I shall argue are ill-conceived, in as much as the essence/accident distinction is itself ill-conceived.

---

9 The essentialist can similarly dispatch Ravnkilde's example (see Ravnkilde, p.61). Ravnkilde cites (A) the necessity of a triangle's being trilateral and (B) the contingency of any particular figure's drawn by Joe being trilateral. Is then the triangle that Joe draws necessarily or contingently trilateral? The essentialist would reply that in so far as (A) holds, then (B) fails to hold: in particular, at least one of the figures drawn by Joe, viz. the triangle he in fact draws, **must** be trilateral.

APPENDIX A

Ravnkilde is typical of a tradition of philosophers beginning with Smullyan who think that the thrust of Quine's misgivings about quantification into modal contexts is that it is illegitimate, because in general quantification into opaque contexts is illegitimate, and modal contexts are opaque. These philosophers have then typically sought to refute the contention that modal contexts are opaque, putting to the side the question of the legitimacy of quantification into opaque contexts. I have of course argued that the impropriety Quine ascribes to quantification into the opaque context of **quotation** is at best the impropriety of a redundant quantifier. But this is not nearly as strong as the claim that such quantification is **logically incoherent**.

Nonetheless I might be faulted for my rather cavalier disregard of the traditional attempts to refute the charge of opacity. We recall that my concern was not so much to refute the charge of opacity as to refute Quine's claim that quantification into modal contexts suffers from an indeterminacy of truth-value. I contended that as long as the essentialist is allowed to interpret the function in a modal context of a definite description with narrow scope as that of a **singular** term, the appearance of indeterminacy dissolves, for since on this interpretation 'the number of planets necessarily is greater than seven' will be materially equivalent to 'nine necessarily is greater than seven', '(Ex)(x necessarily is greater than 7)' will have a determinate truth-value for the essentialist (viz. truth).

Now it might be complained that my strategy depended on interpreting definite descriptions in an unorthodox way, for standardly, i.e. in the classical tradition originating with Russell and Whitehead and continuing through Quine, definite descriptions are not treated as referring expressions/singular terms, rather they are

contextually eliminated according to Russell's theory of descriptions. So if we interpret definite descriptions in the Russellian manner, as Quine would have to, might we not after all find vindicated Quine's contention that modal contexts are opaque? Ought we not at least to consider the Russellian interpretation of Quine's argument for opacity? And correspondingly Russellian refutations of Quine in the manner of Smullyan?

This line of questioning betrays a fundamental confusion: the notion of opacity applies to contextual blanks that are supposed to identify the position of a **singular term**, but since on a Russellian interpretation the position of a definite description does not mark the occurrence of a singular term, there can be no question of describing the position in question as referentially opaque - that would be a category mistake. Now if we stick to proper names construed *contra* Quine as genuine singular terms, then for the essentialist modal contexts have no semblance of opacity, for Quine's criterion of transparency (i.e. **non**-opacity) will be **satisfied** given essentialist intuitions. (E.g. if 'Cicero necessarily is a non-amoeba' is true, so will be 'Tully necessarily is a non-amoeba'). As Neale points out, to endorse the move from 'necFa' to 'necFb' where 'a' and 'b' are coreferential singular terms, is to endorse the view that if an object has a particular property necessarily, then it has that property necessarily however the object is singled out, and thus is to endorse essentialism (Neale, p.140).

In the face of this, one sympathetic to Quine's aversion to QML might change tack, and argue that since there is a Russellian reading of 'the number of the planets necessarily is greater than seven' (henceforth 'S') on which S is **false**, it is important to determine whether S can be derived using standard rules of inference from '9 necessarily is greater than 7' (henceforth 'Q') and 'the number of the planets is nine' (henceforth 'R'), for given our commitment to the standard rules of inference, the

existence of a derivation of **false** S from Q and R would demonstrate the existence of **some** sort of problem with modal contexts, though not the problem of **opacity**. (See Neale, p.136.) To the extent that such a derivation of **false** S existed, **quantification** into modal contexts would inherit the problematic character of modal contexts, so that Quine's exhortation to proscribe QML would be vindicated after all. The question then is whether a Russellian reading of S which is false can indeed be derived from Q and R using standard rules of inference (those found e.g. in Mates' *Elementary Logic).*

There are two possible Russellian readings of S:

S(1) $(\exists x)(y)((Py \equiv y=x) \ \& \ nec(x \rightarrow 7))$

S(2) $nec(\exists x)(y)((Py \equiv y=x) \ \& \ x \rightarrow 7))$

where 'P' is 'is the number of the planets'.

The definite description 'the number of the planets' has wide scope in S(1) and narrow scope in S(2). As Neale shrewdly observes, if S is to be false, as the Quinean would want, then it must be intelligible. But then this rules out S(1) as a possible candidate reading of S for the Quinean, for the intelligibility of S(1) presupposes precisely what Quine wants to discredit, viz. the **intelligibility** of quantification into modal contexts (see Neale, p.136). So S(2) is the sole remaining candidate. And the essentialist would urge that S(2) is false, for they would standardly hold that it is false that **any** of the planets necessarily exists, and thus false that a certain number **had** to be the total number of planets, let alone that it had to exceed seven. But is S(2) **derivable** from Q and R? In fact, it is not. Following Smullyan, Neale points out that **S(1)** is derivable from Q and R, but S(1) is of course of no use here (see Neale, p.137). Thus this tack fails to discredit quantification into modal contexts.

For the essentialist, S(1) is not only intelligible but true. Now as Neale clear-headedly notes, to point to the derivability of S(1) from Q and R (the former can be accepted as true even by an anti-essentialist, albeit as *de dicto* true), in order to establish the **transparency** of modal contexts would be a flagrant *petitio*, for in the derivation of S(1) from Q and R we must allow the substitution of one individual constant for another coreferential constant in a **modal** context, a move which in effect assumes the transparency of modal contexts. The derivation in outline is:

(1)	[1]	9=(ix)(Px)	P
(2)	[2]	□(9>7)	P
(1)	[3]	(∃x)((∀)(Py≡y=x) & x=9)	1, *14.01
(4)	[4]	(∀y)(Py≡y=b) & b=9	P
(4)	[5]	b=9	4 &-ELIM
(2,4)	[6]	□(b>7)	2,5 SUB IDEN
(4)	[7]	(∀y)(Py≡y=b)	4 &-ELIMIN
(2,4)	[8]	(∀y)(Py≡y=b) & □(b>7)	6,7 &-INTRO
(2,4)	[9]	(∃x)((∀y)(Py≡y=x) & □(x>7))	8 EG
(1,2)	[10]	(∃x)((∀y)(Py≡y=x) & □(x>7))	3,4,9 EI

(Neale, p.137)

This deduction clearly assumes the transparency of modal contexts in the derivation of line 6 from 2 and 5 (Neale, p.139). But my concern has not been to establish the **transparency** of modal contexts by pointing to the derivability of S(1) from Q and R.

My point rather is that **S(2)** is not derivable from Q and R. It would appear that as Quine's thoughts progressively clarify over several papers, he comes to recognise a commitment to essentialism as the gravamen of his antipathy to QML. That this might eventually be so is suggested by the consideration that the derivability of S(1) from Q and R depends on the assumption of the transparency of modal contexts (since substitution of coreferentials is made). But then this is, as I have indicated previously, tantamount to an endorsement of essentialism.

CHAPTER 2 THE PROBLEM FOR ESSENTIALISM OF EXISTENCE IN MANY WORLDS

2.1	THE COHERENCY OF THE NOTION OF EXISTENCE IN MANY WORLDS	77
2.1.1	COMMENTS ON CHISHOLM'S FORMULATION	82
2.2	CHISHOLM'S ARGUMENT FROM THE POINT OF VIEW OF MODAL LOGICS	89
2.2.1	THE NECESSITY OF DISTINCTNESS	90
2.2.2	THE NECESSITY OF IDENTITY	93
2.2.2.1	AN ESSENTIALIST REPLY	94
2.2.2.2	ASSESSING THE ESSENTIALIST'S REPLY	96
2.2.3	THE NECESSITY OF DISTINCTNESS	100
2.3	AN ESSENTIALIST REPLY	103
2.3.1	REJECT THE IDENTITY OF INDISCERNIBLES FOR COUNTERFACTUAL POSSIBILITY	103
2.3.2	REJECT THE POSSIBILITY PRINCIPLE	110
2.3.2.1	HAECCEITIES AS UNIQUELY ESSENTIAL	114
2.3.2.2	THE ESSENTIALITY OF UNIQUE WORLD-INDEXED PROPERTIES	120
2.3.3	REJECT THE PRINCIPLE OF THE TRANSITIVITY OF PREDICATIVE POSSIBILITY	123
2.3.4	ARE KRIPKEAN STIPULATIONS A SOLUTION?	126
2.4	LOOKING AHEAD TO CHAPTER 3	128

CHAPTER 2: THE PROBLEM FOR ESSENTIALISM OF EXISTENCE IN MANY WORLDS

2.1 THE COHERENCY OF THE NOTION OF POSSIBLY HAVING BEEN OTHERWISE

We recall from section 1.2 that the essentialist's point of departure is the widespread pre-analytic intuition that some things could have been otherwise in some particular respect. Take for example some actual object a: a is F, but perhaps - pre-analytically at any rate - it could have been not F. So we could say that the circumstances of a's being not F are **possible though non-actual circumstances for a,** for they are circumstances in which it is **a** that is not F. Of course the customary practice for philosophers who presently endorse this pre-analytic intuition is - to my mind unfortunately - to refer to this situation as one in which there is a **possible world** in which a exists and is not F. According to this idiom: (1) a exists in a possible world w just where the obtaining of certain circumstances which comprise w, would have included/would include the obtaining of the existence of a; (2) a is F in a possible world w just where a would have been/would be/is F if the circumstances comprising w had obtained/were to obtain/do obtain. Since a is contingently F just where, though it is F, it could have been/could be/is not F - i.e. the circumstances of a's not being F are possible though non-actual circumstances for a - then, in the idiom of possible worlds, a is contingently F just where there is a possible world in which 'Fa' is true and a possible world in which it is false. Similarly, since a is essentially F just where no circumstances of a's not being F are possible circumstances for a - i.e. circumstances in which it is **a** that is not F - a is essentially F just where there is no possible world in which a exists and is not F; alternatively, a is essentially F just where it is F in every possible world in which it exists.

So in the idiom of possible worlds, a's having a property contingently involves the notion of a's existence in more than one possible world: a is contingently F just where there is a possible world, viz. the actual one, in which a exists and is F and there is a non-actual world in which a exists where it is not F; therefore a exists in at least two possible worlds if it is contingently F.

Now according to an objection originating with Chisholm in his *locus classicus* 'Identity through possible worlds: Some Questions', the notion of an object's existing in multiple possible worlds, which is the notion that an object could have been otherwise than it actually was, leads to contradiction. Since Chisholm's argument applies to any actual object, it follows that if sound, it renders the notion of contingent properties in general incoherent, and thus renders non-existent the alleged distinction between contingent and necessary properties fundamental to essentialism. So much the worse for essentialism. The thrust of Chisholm's anti-essentialist argument is most perspicuously brought out in the following reconstruction.

Take two **actual** and **contemporaneous** objects a and b. By hypothesis they are distinct, originate **simultaneously** - perhaps at a great spatial distance from each other, perhaps not - and perish **simultaneously**. (The stipulation of contemporaneity will be explained in 2.1.1.) Both objects will have the same number of properties, since for any property, either it or its complement (but not both) is instantiated by both objects. Let '$U_{i=1}F_i$' represent the conjunction 'F_1 & F_2 & F_3 & ...' of all of the predicates corresponding to a's properties. Let '$U_{i=1}G_i$' represent the conjunction 'G_1 & G_2 &G_3 & ...' of all the predicates corresponding to b's properties. I leave it open whether a and b in fact share any properties other than those of originating at the same time and perishing at the same time. Clearly mass-produced screws will share many properties; while some organism or object

which originated at the same time I did may not share much with me. Moreover, a and b will each have an **indenumerably infinite** number of properties, for their properties are meant to be **temporally indexed**, and any temporal interval will have the cardinality of the set of the real numbers. Thus e.g. a might be *inter alia* F_{10} and F_{15} , where 'F_{10}' is 'green-at-t_{100}', and 'F_{15}'is 'red-at-t_{300}'. (Strictly speaking this raises a problem of notation: the subscript 'i' in 'U_i F_i' ranges over integers; but since there are non-denumerably many properties involved then a **list** of them will have to use subscripts that are not integers - F_1 , F_2 , ..., F_ω , $F_{\omega+1}$, ... For simplicity I will not adapt the notation to this point about non-denumerability. Nor will I take up the point that a super-denumerable **language** will be required.) Now imagine an object that differs from a at t in only one respect, namely it is G_1 where a is F_1, (assuming now that $F_1 \neq G_1$), so that this object is G_1 & $U_{i=2}$ F_i i.e. it is G_1 & F_2& F_3& F_4& ..., while a is $U_{i=1}$ F_i , i.e. it is F_1 & F_2 & F_3 & F_4 & .... Chisholm asks whether the object in question is such that **a** could have **had** its properties. An affirmative answer would appear to appeal to the general principle, which I shall call 'the possibility principle', that an object x could have had exactly the genuine properties of an object which differs from x only in the circumstance that one of x's genuine properties is replaced by another not possessed by x. In all probability the object that differs in this way from x will be non-actual given that a is actual, for even artifacts cast from the same mould would presumably differ in more than just one respect. (I assume for ease of exposition that where x is characterised both by a positive determinate that falls under a recognised determinable and by a corresponding negative determinate - thus for example suppose that x is red and thus non-green - then a replacement in x of red with green also involves as the **same** replacement, and **not** a separate replacement, a replacement in x of **non-green** with green. Otherwise x will be both green and non-green after one replacement, namely the replacement of red with green. Similarly, if someone were to insist that if x

counts F among its properties then it also counts not-not-F among its properties, then I would stipulate that in replacing F with not-F we also automatically replace not-not-F with not-F as part of the same replacement. Otherwise the possibility principle holds on pain of rendering objects such that they could have had inconsistent properties at the one time. But these are all tiresome niceties mentioned here merely to forestall if necessary the cavils of an uncharitable pedant.)

So by the possibility principle, a which is $U_{i=1}F_i$, could have had exactly the properties of an object that is G_1 & $U_{i=2}F_i$. . And again by the possibility principle the latter object could in turn have had the properties of an object that is G_1 & G_2 & $U_{i=3}$ F_i (assuming now that $F_2 \neq G_2$). At this point the essentialist might want to invoke a further general principle to the effect that if an object a that is R (i.e. is predicatively R rather than is identical to R) could have had exactly the properties of an object that is predicatively S, then if an object that is predicatively S could in turn have had exactly the properties of an object that is predicatively T, then a could also have had exactly the properties of an object that is predicatively T. Let us call this 'the principle of the transitivity of predicative possibility' or 'the transitivity principle' for short. (The insertion of the adverb 'predicatively' throughout the statement of this principle is to underscore that it is a principle of the transitivity of **predicative possibility** as opposed to the transitivity of **identity**, for more of which contrast see below.)

It follows by the transitivity principle that if a could have been G_1 & $U_{i=2}$ F_i then a could have been G_1 & G_2 & $U_{i=3}$ F_i . Clearly, we can proceed in the manner above, i.e. with iterated applications of both the possibility principle and the transitivity principle, to construct a sorites of hypothetical premisses with the conclusion that if a could have been G_1 & $U_{i=2}$ F_i then a could have been $U_{i=1}$ G_i. And since by the possibility principle a could have been G_1 & $U_{i=2}$ F_i it follows that a could have been

$U_{i=1}$ G_i. . The proceeding discussion can be rerun for b - which is $U_{i=1}$ G_i - to yield the conclusion that b could have been $U_{i=1}$ F_i . It follows that a could have had the properties of b. Let us call the principle that if **actual** a could have had all of **actual** b's properties, then a could have been b, 'the Principle of the Identity of Indiscernibles in Cases of Counterfactual Possibility' or 'PIOICP'. If PIOICP is, as I shall argue, a valid principle, then a could have been identical with b, for the former is, and the latter could have been, $U_{i=1}$ F_i ; that is, a could have had all of b's properties, but then - as I shall argue - a could have been b. Similarly b could have been identical with a, for the former is, and the latter could have been, $U_{i=1}$ G_i ; that is, b could have had all of a's properties, but then - as I shall argue - b could have been a.

What this amounts to is the prospect that given an object's existence in multiple worlds, then where a and b are two actual contemporaneous objects, a could have been b and b could have been a. But this is **deeply** puzzling, indeed incoherent; after all what does it mean to say that a could have been b and b could have been a? A is not b; b is not a. How could a have been something that is not a and yet still have been a? Similarly, how could b have been something that is not b and yet still have been b? Moreover, if a could have been b, why would it have been a who is b, as opposed to someone or something other than a? In other words, what is it about the situation, in which a allegedly has all of b's properties, that identifies the bearer of all of b's properties as **a** and not something other than a, when by hypothesis a has relinquished all of its actual properties? Similarly, if b could have been a, why would it have been be b who is a, as opposed to someone or something other than b? To respond that merely by stipulative fiat we secure the identity of the bearer of all of b's properties as that of a, is to shift the subject from metaphysical possibility to conceptual possibility. If possibility is just about the stipulations involved in our concepts then of course anything could have been anything merely by rigging our

concepts or stipulations in the appropriate way. But this is not philosophically interesting in the way metaphysical possibility is supposed to be. Metaphysical possibility is supposed to be a function of the the way the world is and operates. The limits it sets on the shape of that world are supposed to be objectively determined in a manner independent of human interests and concerns. So to **stipulate** baldly and glibly that a bearer of all of b's actual properties is a is in effect to whistle in the metaphysical dark.

Given that essentialism turns on the claim that things could have been otherwise, and given that this claim is committed to recognising as coherent what is incoherent, it follows that the claim that things could have been otherwise, in the metaphysical sense of 'could', is incoherent, and thus that essentialism is incoherent. The distinction alleged by essentialists to obtain between essential and accidental properties is incoherent. This is not to say that as it turns out individuals have essential properties only or accidental properties only; rather it is to say that any property an individual has is better described as neither essential nor accidental. The dichotomy is unintelligible.

2.1.1 COMMENTS ON CHISHOLM'S FORMULATION

Note that my formulation of Chisholm's argument has been influenced by Forbes' adaptation of it to the narrower question of the essentiality of an artifact's material composition in his *The Metaphysics of Modality*. Chisholm of course was originally concerned not just with the question whether an artifact could have arisen from a slightly different material origin, but with the implications of the wider notion that an object - artifact or non-artifact - could have been different from how it actually was in any respect.

Now in examining this question Chisholm spoke of an object-in-one-possible-world's being **identical** with an object-in-another-possible-world. My formulation by contrast states that on the supposition that an actual object a differs from some non-actual object a_1 by just one property, the essentialist would conclude by the possibility principle that a could have **had** all the properties of a_1. I refrained from saying that the essentialist would conclude that a could have **been** a_1. Chisholm for his part says that in that situation, a, which exists in the actual world, is **identical** with a_1, which exists in a non-actual possible world w_1. (Note the indicative 'is' in 'is identical with'.)

Now by hypothesis a_1 in turn differs by just one property from another non-actual object a_2 existing in w_2. Thus, by Chisholm's lights, a_1 is identical with a_2. And a_2 differs by just one property from a_3 which exists in w_3, so a_2 is identical with a_3, etc., ... and finally a_{n-1} which exists in w_{n-1}, differs by just one property from a_n which exists in w_n, so that a_{n-1} is identical with a_n. But a_n is indiscernible from b, which by hypothesis exists in the actual world and is distinct from a. So Chisholm concludes that a_n is b, rather than admit the possibility that the actual world could have been just as it was except for the circumstance that a_n, which is indiscernible from b, exists in place of b as a distinct entity. In so doing Chisholm appears to follow a Leibnizian line of reasoning; specifically, he appeals to a form of the Principle of Sufficient Reason: God would have had no sufficient reason for choosing a world containing a_n rather than a world containing b (Chisholm 1967, p.83). So $a=a_1=a_2=a_3= \ldots =a_{n-1}=a_n=b$. Thus, by the transitivity of identity, a=b, which contradicts the assumption that a and b are distinct. (cf. Loux, p.38)

Clearly, I have followed Chisholm in my use of **non-actual** individuals in my reconstruction of his argument: I said that an actual object a could have had all the properties of an object a_1 from which it differs by just one property; but then in all

probability a_1 would be **non-actual.** After all, actual artifacts cast from the same mould typically differ in more than just one respect. But unlike Chisholm I did not speak of a's being **identical** with a_1, and unlike Chisholm I did not use the transitivity of **identity**. Instead, I followed Forbes in my use of a transitivity principle, which was not however the transitivity of identity, but of **predicative** possibility. Recall: if a could have been F, and something which is F could have been G, then a could have been G. Note that I did not say that a could have been **the object** which is G. My conclusion was merely that a could have had all the **properties** of the object which is G. Contrast this with Chisholm's approach: if a in the actual world is a_1, and a_1 is a_2, than a **is** a_2.

The reason why I diverged from Chisholm's approach is that I find suspect a **literal** construal of talk of an object-in-one-possible-world's being identical with an object-in-another-possible-world. That is to say that I find suspect the notion that an object in w_i is (**indicative** 'is') literally identical with an object in w_j. The entity that I am exists in the actual world; it does not co-exist with or pre-exist or post-exist other alternative I's, not even when these other I's are said to be in logical as opposed to physical space - whatever that means. The I I am is real whereas the other I's are not. To maintain otherwise is to have lost that robustly Russellian sense of reality. Thus I who am real cannot be identical with an I which is not. They are not the same. By contrast, the boy who I was is identical with the man who I am. They are the same. But then they are both located in the same physical space, the actual space. In fact, any plausibility attaching to the notion that an object is identical with a non-actual counterpart (which I here understand in an ordinary, non-technical, i.e. non-Lewisian sense of 'counterpart'), probably derives from a confused conception of this counterpart as being equally located in the same physical space, so that just as continuants survive minor physical changes, so do they survive minor modal changes: i.e. just as the person I was last night is identical

with the person I am this morning, the object I actually am is identical with the object I could have been had I been born, say, five minutes earlier. So an essentialist who shared my misgivings about identity across worlds would agree that Chisholm's argument, as he formulates it, does not get off the ground, in as much as it appeals to an unacceptable notion of identity across worlds, a notion the essentialist in question will aver is one to which they need not be committed.

Perhaps Chisholm has not been misled by a reification of non-actual worlds - such reification being to my mind one of the dangers attending the usage of possible worlds - and thus did not mean that an actual object a is numerically identical with a non-actual object a_1. Perhaps all he meant was that a **could have** been (identical with) a_1. But what would this mean? It seems hardly more intelligible to say that a **could have** been numerically identical with a_1 than to say that a **is** numerically identical with a_1. Indeed I have argued that it is just this sort of puzzling, indeed incoherent thesis to which the essentialist is committed (in particular, the thesis that a couId have been b). In that case Chisholm's argument against essentialism doesn't even get started, for by my own admission its premises are unintelligible. Faced with this difficulty and my obligation to the esentialist to formulate the premises of Chisholm's argument in a manner acceptable to them, I thus offer something like the following as an explication of the claim that a could have been a_1, and a_1 could have been a_2, etc.: in a replay of the history of the universe, everything proceeds as it actually does until a certain time t, after which time the course of events diverges from actuality sufficiently for a to have all the properties of a_1; similarly for a1 with regard to a_2, etc. (Alternatively, the universe could have forked off into a non-actual pathway such that a had all the properties of a_1 etc.). This then in effect brings us to my version of Chisholm's argument with its talk of a's possibly having had the properties of a_2, rather than being a_2, etc.

But then the essentialist might protest that the problematic conclusion of Chisholm's argument - viz. that actual a could have been identical with actual and distinct b - does not follow. For strictly speaking, given the latest explanation of 'a could have been a_1, a_1 could have been a_2, ... , a_{n-1} could have been a_n (which is b)', all my reconstruction licenses is merely the conclusion that a could have had all the properties of b. So how does it follow from this that a could **literally** have been b, which I urged is the problematic conclusion to which the essentialist is committed? Certainly it does not follow as a corollary. Nonetheless, I believe there is a good ground for inferring it. For quite simply if a could have had all of b's properties, why would it be **a** that could have had b's properties rather than something other than a? What in other words secures the identity of a **bearer** of all of b's properties as that of a? And in general, what secures the identity of an arbitrary object x that is said to bear all of b's properties as indeed that of x? Only the vague insinuation that what individuates x is a propertyless substratum peculiar to x, such that it is this substratum that acquires all of b's properties, could to my mind explain the temptation to identify the bearer in question of all of b's properties as x. But I do not accept such an account of individuation; propertyless substrata or bare particulars do not individuate, since they do not exist. And even if we suppose that they did exist, how would they in fact secure the identity of their correlative individuals when they are after all bare, i.e. propertyless? Failing a satisfactory answer to the question as to what would secure the identity of a bearer of all of b's properties as a non-b, I conclude that a bearer of all of b's properties is nothing other than b. A bearer of all of b's properties must be b. But in that case **a** could have been b, for it could have been a bearer of all of b's properties. And it is just this claim of identity which I suggest is puzzling, indeed incoherent, and to which the essentialist's intuition that an object could have been otherwise commits them.

(Incidentally, it follows by parity of reasoning that a could have been literally

identical with a_1; that a_1 could have been literally identical with a_2; etc. So does it follow that I could have followed Chisholm after all in his use of the transitivity of identity, instead of invoking as I did the transitivity of predicative possibility? That depends on whether a statement of the form 'if a could have been a_1, and a_1 could have been a_2, then a could have been a_2', counts as a statement of the transitivity of identity, for standardly the latter is expressed in the indicative as opposed to the subjunctive.)

Another respect in which I have diverged from Chisholm's presentation of his argument is that I have temporally indexed a's and b's properties; for his part Chisholm writes as if a and b, in his case Adam and Noah, are immutable, and so reference to the time of a property's possession becomes redundant. Of course concrete objects do change; hence in comparing two objects that differ with respect to the possession of a temporally **non**-indexed property it is necessary to specify at what time this difference obtains, for at another time the objects in question may not differ with regard to that property.

A related consideration which necessitated yet another divergence from Chisholm's presentation is that Chisholm's reductio works better if we stipulate, as I have, that a and b be **contemporaneous**. After all the essentialist could justifiably protest that Chisholm's argument could hardly get under way if it requires that a could have had the property of being F at a certain time t, when a does not originate until well after t. To explain myself: take Noah for instance. Consider any property that Adam has but Noah lacks. Since Noah does not come into existence until after Adam perishes then any of Adam's temporally indexed properties will fill the bill. Now consider the circumstance in which an anonymous object is just like Noah but for the replacement of one of Noah's properties with one of Adam's temporally indexed properties. The possibility principle guarantees the coherence of this hypothesis.

But how could the description in question represent a possible state of affairs, for another property of the anonymous object under discussion is that it comes into existence only **after** Adam has perished? Furthermore, according to the possibility principle Noah could have had **all** of the properties of this anonymous object. But this would mean that Noah could have had a temporally indexed property of Adam. So according to the possibility principle Noah could have had a certain property even before **he** existed. These difficulties are most embarrassing for a proposed *reductio* of essentialism that trades on the possibility principle. Clearly an essentialist would be within their rights in disavowing such a principle. To remedy these difficulties for the anti-essentialist I have stipulated that a and b be contemporaneous.

At this point the essentialist might protest that even this remedy will not do, for one can always ask how a, which is continually green, could have had **all** of the properties of an anonymous object that differs from it by one property, where the property is that of being red at a certain time t. In other words could a, which is actually green at **all** times, have been such that while otherwise green it is instantaneously red just once? The point would be well taken if the notion of 'could have' involved were that of causal possibility. Presumably actual things causally interact - at least at the macroscopic level - in such a way that a process does not issue in an end state which differs in just one respect from the initial state. If a could (causally could) have been red at t then presumably this was so only if a was also different in further respects. So the possibility principle would probably not apply realistically to causal possibility. Moreover we don't take it to be causally possible - at the macroscopic level anyway - that things could (causally could) be red instantaneously and otherwise green. So again the possibility principle would fail as a representation of causal possibilities. But the point that must be emphasised here is that we are not dealing with causal possibility but metaphysical possibility. And

the essentialist would have to grant that, by their own lights, it is metaphysically possible that an object be red instantaneously and otherwise green. Admittedly it is highly unlikely if not causally impossible. But for a strict causal determinist anything non-actual which the essentialist endorses as nonetheless metaphysically possible will be causally impossible. So if we open the doors to some indeterminism, we then cannot without arbitariness close them to the metaphysical possibility of an object's being red instantaneously and otherwise green. But then the possibility principle as applied to metaphysical possibility will have to be allowed by the essentialist. It is not my brief to defend the possibility principle or to explain just how it is that an object could metaphysically have been red instantaneously and otherwise green. Rather, given that I have established that the essentialist ought to accept the possibility principle, I have laid a plank necessary to support a reductio of essentialism.

2.2 CHISHOLM'S ARGUMENT FROM THE POINT OF VIEW OF MODAL LOGICS

From a formal point of view, what Chisholm's argument shows is that if an object a could have been otherwise - or expressed in term as of possible worlds: if a exists in more than one possible world - then while a is not actually identical with another object b, a could have been identical with b. Now clearly this contradicts a theorem of S5 lower predicate calculus with identity, viz. '$(x)(y)[x\neq y \rightarrow \text{nec}x\neq y)]$' - call this 'the necessity of distinctness thesis', or 'LNI'. Thus S5, which endorses LNI, will have to be abandoned if one believes that things could have been otherwise. But even systems of modal logic which do not contain LNI as a theorem will have to be abandoned. This is because the claim that an object **could have** been another object is incoherent. To repeat: what does it even mean to say that a could have been b and b could have been a? A is not b; b is nit a. How could a have been something

that is not a and yet still have been a? Similarly, how could b have been something that is not b and yet still have been b? The upshot is that we should reject as incoherent the idea that an object could have been otherwise. Thus we should also reject as incoherent the alleged distinction between metaphysically - as opposed to merely conceptually - essential properties and metaphysically accidental properties, which distinction the idea that an object could have been otherwise informs. The distinction between essential and accidental does not so much collapse into the essential as into the incoherent.

2.2.1 THE NECESSITY OF DISTINCTNESS

Taken in terms of possible worlds, LNI states that in general, if an object a is distinct from an object b, then it is **necessarily** true that a is distinct from b, i.e. it is true in all possible worlds that a is distinct from b. So there will not be a possible world in which a is identical with b. So if a is actually distinct from b, then a could not have been identical with b. In particular, if I am distinct from my contemporary, then I could not have been identical with him. For its part LNI depends proof-theoretically on both the S5 theorem schema 'p $\rightarrow \Box\Diamond$p', the so-called Brouwerian axiom, and a theorem of S5, viz. '(x)(y)[Ixy $\rightarrow$ (necIx)y)]', where 'I' is the identity predicate. This theorem is the necessity of identity thesis or LI[10]. According to LI

---

10 Standard presentations of LI render it thus: (x)(y)(x=y $\rightarrow$ $\Box$x=y). This is to blur the distinction between `(x)(y)(Ixy $\rightarrow$ nec(Ixy))' and `(x)(y)(Ixy $\rightarrow$ (necIx)y)'. It is the latter which is important for our concerns here. Similarly the formalisation of LNI that interests us here is `(x)(y)(-Ixy $\rightarrow$ (nec-Ix)y)' rather than `...$\rightarrow$ nec(-Ixy))'. Nonetheless on the semantics

if an object designated 'a' is identical with an object designated 'b', then it is **necessarily** true that a is identical with b. That LNI depends on 'p → □◊p' and LI is shown in the following adaptatation of a proof by Wiggins in S5:

1	1	a≠b	P
	2	p → □◊p	S5 theorem schema
	3	a≠b → □◊(a≠b)	2
1	4	□◊(a≠b)	1,3 mod pon
1	5	-◊-◊(a≠b)	4 □↔–◊–
6	6	-□(a≠b)	P for reductio
6	7	◊(a=b)	6 ↔-□-
	8	a=b ↔ (□Ia)b	identity theorem
	9	a=b ↔ □(a=b)	8 equiv of '(necIx)y' and 'nec(Ixy)' (see n10)
6	10	◊□ (a=b)	7,9 substi log equiv
6	11	◊-◊(a≠b)	10 □↔-◊-
1,6	12	-◊-◊(a≠b) & ◊-◊(a≠b)	5,11 conjunc
1	13	--□(a≠b)	6,12 RAA
1	14	□(a≠b)	13 double neg

outlined in chapter 1 it follows that for LI the disambiguated sentences are logically equivalent. Similarly for LNI.

15	$a \neq b \rightarrow \Box(a \neq b)$	1,14 conditional
16	$a \neq b \leftrightarrow$ (nec-Ia)b	equiv of 'nec(-Iab)' and '(nec-Ia)b' (see n10)

(Wiggins 1980, p.217)

The provenance of line 3 is the Brouwerian axiom'$p \rightarrow \Box\Diamond p$' in line 2, which sentence in turn is derived in S5 from the characteristic S5 axiom '$\Diamond p \rightarrow \Box\Diamond p$' and the S5 (T) theorem '$p \rightarrow \Diamond p$'.(see Hughes and Cresswell, p.57). The derivation of line 10 from 7 and 9 requires that 'a=b' be logically equivalent to '$\Box$(a=b)', which in turn requires that LI be true, i.e. that '$a=b \rightarrow$ (necIa)b' be a theorem of S5. LI is of course a theorem of S5 (and also of S4 and T). (see next section). The upshot of Chisholm's argument is that S5 cannot be the correct modal logic to express the notion of counterfactual possibilities and the essential/accidental distinction which this notion informs.

An essentialist committed to S5 might try to modify standard S5 lower predicate calculus with identity, by restricting one of the identity axioms so as to obtain an S5 with **contingent** identity, thus blocking the derivation of LI (see Hughes & Cresswell 1972, p.195) and thereby of LNI. They may even try to motivate this modification with the thought that some statements of identity are after all contingently true. But this position is far from uncontroversial even among essentialists; furthermore, the S5 essentialist who espouses contingent identity will have to oppose the standard Barcan-Kripke proof of LI. These are issues which I shall now address.

2.2.2 THE NECESSITY OF IDENTITY

A Barcan-Kripke proof of LI valid in T, S4 and S5 tends to run along something like the following lines:

1) if an object x is identical to an object y then every property of x is a property of y and *vice versa* (Indiscernibility of Identicals)

2) every object x has the property of being **necessarily** identical to x i.e. **(x)(necIx)x ,** where 'I' is the identity predicate

Therefore:

3) if x is identical to y then the property of being necessarily identical to x is a property of y

(from 1 and 2)

i.e. **LI**: if x=y then (necIx)y.

Note that the property introduced in (2) is not the property of self-identity, which if it existed would be common to all individuals, but rather the property - where x is a, say - of having **necessarily** the **haecceity** of a, where the haecceity of a is the property of being identical to a. (Thus the haecceity of b would be the property of being identical to b). Now the essentialist will standardly urge that having necessarily the haecceity of x (being necessarily identical to x) is a genuine property that is unique to the x in question, just as x's haecceity is unique to it. So given LI the proof of LNI in S5 can proceed. And the problem for the essentialist then becomes that of harbouring the inconsistency that I could have been identical to my contemporary, though I am actually distinct from him, while LNI counsels the very opposite conclusion. Moreover, given LI, it follows that if x is identical to y, then y

could not have been distinct from x., for y is not only identical to x but necessarily so. Thus given that I am identical with myself, I could not have been identical to my contemporary, for in being my contemporary I would have had to be something other than I, i.e something distinct from me. And yet by Chisholm's argument the essentialist is committed to this very possibility and thus to a further inconsistency, viz.the conjunction of this possibility with LI.

To dispose of these two *ad hominem* difficulties, the essentialist might want to repudiate the assumption in (2) that having necessarily the haecceity of x is a **genuine** property of x. But if this is not a genuine property of x, then why should its haecceity be either? After all, essentialists standardly urge that where F is an essential property of x, being F essentially or necessarily counts as a further property of x. And so it should if the essence/accident distinction is a sound one. So if being necessarily identical to x is not a genuine property of x, then neither should being identical to x, i.e. x's haecceity, be. a genuine property of x. But then - to anticipate the discussion ahead - the essentialist will be powerless to resist the conclusion of Chisholm's argument that if a could have had one of b's properties, then it could have had all of b's properties, and given this conclusion , I have argued that a could have **been** b, *per impossibile*. The essentialist needs haecceities to thwart this conclusion. See 2.3.2 - 2.3.2.1.

2.2.2.1 AN ESSENTIALIST REPLY

Of course an essentialist whose patience has been overtaxed might well want to reply at this point, perhaps with barely restrained irritation, that it has all along been simply wrongheaded of **me,** to describe as **incoherent** the alleged possibility that a could have been b. Pointing to arguments by Gibbard and Lewis they will insist that

in cases of material composition the matter which composes an object, while identical with it, could have been distinct from it. And since the matter could have been distinct from it, the matter could have been non-identical with it, and thus from itself, for the matter is the object which it composes So the matter, while identical with the object it composes, is not necessarily identical with this object and thus not necessarily identical with itself.

Gibbard's argument involves a statue which is formed by combining two portions of pre-moulded clay. Since clay dries out over time, and Gibbard wants to say that the statue is identical with the **same** lump of clay at all times, let us say that the statue is formed from the combination of two pieces of pre-moulded **plasticine.** Plasticine doesn't dry out. Suppose that the combined lump of plasticine is destroyed at some later time, perhaps vaporised. Then clearly the statue will also have been destroyed. So the lump of plasticine and the statue will have been contemporaneous throughout their respective histories. Gibbard concludes that the two are identical.

> They began at the same time, ... they had the same shape, location, color, and so forth at each instant in their history; everything that happened to one happened to the other; and the act that destroyed the one destroyed the other. If the statue is an entity over and above the ... [specially moulded lump of plasticine that constitutes it], then statues seem to take on a ghostly air. (Gibbard 1975, p.191)

Sainsbury reinforces Gibbard's conclusion with the questions 'How could two things occupy exactly the same place at exactly the same times? If ... [both weigh two pounds] why do not the two together weigh four pounds?' (Sainsbury 1995, p.94.) But if the lump of plasticine is identical with the statue, it is not necessarily so, so Gibbard urges. For if things could have been otherwise, then the statue could have been destroyed by pressure sufficient to distort its shape (form) beyond recognition without an accompanying loss in the material integrity of the lump which constitutes it. But then the lump could have **survived** the destruction of the statue. Hence by the Indiscernibility of Identicals the lump would not have been

identical with the statue. Thus the lump, while identical with the statue, is only **contingently** so.

Lewis' argument (Lewis 1986, pp.252-253) is the same in essentials as Gibbard's: in a factory a lump of plastic is synthesised in a bowl-shaped mould. Thus the lump of plastic which constitutes the bowl comes into existence simultaneously with the bowl. Since, as Sainsbury comments, at every instant both or neither of the bowl and the lump of plastic exist, and moreover in the same place, with the same weight, etc., it is reasonable to affirm the identity of the lump of plastic with the bowl (Sainsbury 1991, p.271).[11] Now according to Lewis it could have been the case that the original lump of plastic was synthesised in a basket-shaped mould while a second lump of plastic was synthesised in the bowl mould. Lewis thinks that under these conditions the original lump of plastic and the original bowl would have both existed but as distinct objects, the latter as an object constituted by the second lump of plastic. So the original plastic could have been distinct from the bowl, which in turn could have been distinct from the original plastic.

2.2.2.2 ASSESSING THE ESSENTIALIST'S REPLY

How convincing are these arguments? It would be question-begging of me as an anti-essentialist to respond that these arguments should be rejected in as much as they presuppose something incoherent, viz. that things could have been otherwise,

---

11 If these sorts of considerations lead Gibbard and Lewis to identify the matter with the object it constitutes, then they ought equally to lead them to endorse PIOICP. For among the properties of an object b which object a could have had will be b's material composition, trajectory through space over time, etc.

because the reason I urge for the incoherence of this presupposition just is that it leads to what I want to deny and the essentialist wants to affirm, viz. that an object could have been distinct from itself. Instead my strategy will be to grant the essentialist for the sake of argument that things could have been otherwise. This will be with a view to determining whether Gibbard's and Lewis' thought experiments need be construed in the manner they would urge they should.

Note that essentialists like Kripke and Salmon would not concede the force of these arguments. They would urge that an object's origin is essential to its identity, so that in the case of Lewis' plastics factory, the bowl made from the second lump of plastic instead of the original lump would not count as the original bowl, thus stalemating this alleged counterexample to the necessity of identity thesis. (I shall pursue the issue of the essentiality of origins in chapter 3.) But what of the other alleged counterexamples, viz. that the original lump of plastic could have been synthesised in a basket mould instead of a bowl moul; and that the lump of plasticine could have survived deformation even though the statue it constitutes couldn't have?

As for the former case, it is not clear that an essentialist must grant that the original lump of plastic is identical with the original bowl. In fact Della Rocca's discussion in Della Rocca 1996 attacks what he takes to be the general and erroneous perception among anti-full--essentialists that full-essentialists are **committed** to the **non-identity** of the bowl with the lump. (His actual discussion concerns a statue and a lump.) As an essentialist Della Rocca himself displays an ambiguous stance with regard to the issue of the identity or non-identity of the bowl and the lump. On the one hand he appears to be agnostic, stressing that he is not claiming identity (Della Rocca 1996, p.198) and granting that there may be, for all he knows, an argument for non-identity which is non-circular, unlike the one he discusses in his

paper. (Della Rocca 1996, p.201.) On the other hand he describes as obvious the fact that the lump and the bowl are in some sense the same (Della Rocca 1996, p.189); furthermore he regards **seriously** the view of some philosophers that the lump and bowl are clearly identical (Della Rocca 1996, p.201): he regards it seriously enough to want to undercut the claim that full-essentialists are **committed** to the non-identity of lump and bowl. (This he does by showing up the circularity of an argument based on the Indiscernibility of Identicals and the further premisses that the bowl is essentially a bowl while the lump is not essentially a bowl. In so far as the latter two premises are adduced by the essentialist as bare claims of allegedly uncontroversial intuition, the argument is circular, for to have a reason to believe these two premisses requires that one have already settled the issue of identity or non-identity of lump and bowl. (Della Rocca 1996, pp.196-199.))

For his part Wiggins is one essentialist who would unambiguously urge that the lump and bowl are not identical. The relation between them on his view is not identity but constitution: the plastic constitutes the bowl; and is distinct from it; that is why it could survive the destruction through deformation of the statue. So if the plastic is distinct from the bowl it constitutes, there is no question of describing the situation as one in which the plastic is only contingently identical with the bowl. Similarly, in the latter case, the plasticine constitutes the statue rather than being identical with it; so again there is no question of identity, let alone contingent identity.

But what is this relation of constitution? Lowe confesses that he has no settled view on a reductive analysis of this notion (recall Gibbard's observation that statues as objects over and above lumps take on a ghostly air). Clearly a constitutes b only if their respective spatial positions coincide. Furthemore the relation is asymmetric: if a constitutes b then b does not constitute a. Lowe even suggests that some or all of

the properties of b supervene on those of a (Lowe 1989, p.81). But even in default of such an analysis we can grasp intuitively the distinction between an object and its constituent matter. Moreover, a separate consideration reinforces the viability of this distinction. For armed with this distinction we can resolve the Heraclitean paradox of not being able to step into the same river twice at the same spot. The solution of the paradox is that we can step into the same river twice but not into the same water. The water merely constitutes the river at the spot where we step in. So a change in water does not necessitate a change in the river.

As for the objection that the volume occupied by a distinctively shaped mass of two kilograms weighs only two kilograms and not four, this can be disarmed. As Sainsbury points out, this problem remains even if we affirm the contingency of identity, so it provides no reason for affirming it.

> Consider a case in which ... a statue and a lump are distinct, for example because the lump comes into existence before the statue. None the less, [Gibbard and Lewis will have to admit that] there is a period of time during which two things occupy exactly the same places at the same time; and during which each has a weight and the weight of the two together is the same as the weight of each (Sainsbury 1995, p.94)

Lewis might want to respond to Sainsbury's example by adopting a view of objects as four-dimensional space-time worms rather than as continuants. On this view, as a four-dimensional object the statue in Sainsbury's example shares with the lump a spatio-temporal part, viz. that region of space-time described in the continuant idiom as a region of space occupied by two objects over the same period. Now there is only **one** spatio-temporal part in that delimited region of space-time, not **two** coincident spatio-temporal parts, so there is no problem of two kilos rather than four. So Lewis could urge that Sainsbury's example doesn't present the same problem when reformulated in four dimensional language. But I disagree. For the question now becomes 'how is it that two distinct space-time worms can share the same spatio-temporal part?' If it is problematic how two continuants can coincide

spatially over a period of time, whether for the duration of their existence or only a part of it, then surely it is equally problematic how two space-time worms can share the same spatio-temporal part. If, as Lewis complains, it reeks of double counting to distinguish the statue from its plasticine or the bowl from its plastic (Lewis p.252), then it equally ought to smell when we distinguish the 4D sub-region of Sainsbury's lump, i.e. the statue, from the larger 4D region, i.e. the lump, of which it is a part. So if double counting is a problem it remains one under a 4D interpretation of objects. (In any case, it is not clear that in compiling an inventory of how many different objects there are we do not count the parts of a gross object, as well as the object, their parts and so on down to the microscopic or subatomic level.) As for there being only two kilos rather than four in the one spot, that is not a peculiarity of constitution. Certainly the problem is not evaded on a 4D response to Sainsbury's example: the combined weight of the 4D lump and the 4D statue is not the weight of the lump plus the weight of the statue, as we would expect given that the lump and statue are distinct, but just the weight of the 4D lump. The upshot is that an essentialist need not accept Gibbard's or Lewis' decription of their respective examples as one in which the constitutive matter is **identical** with the object it constitutes, and hence the case for contingent identity does not get off the ground.

2.2.3 THE NECESSITY OF DISTINCTNESS

So given the difficulties which attend an attempt by the essentialist committed to S5 to adopt contingent identity, the essentialist might choose instead to deny the characteristic S5 axiom '$\Diamond p \rightarrow \Box\Diamond p$', which is thus to reject S5 altogether. Now while this move will effectively block the derivation of the necessity of distinctness thesis LNI in systems other than S5, such as S4 or T, we note that given essentialist intuitions about possibility and necessity, it is 'at least arguable that intuitively [at

least for an essentialist] LI and LNI stand or fall together, and that if a satisfactory modal system is to contain LI it should contain LNI as well' (Hughes & Cresswell 1972, p.192). So since standard T and standard S4 both contain LI, they should also both contain LNI. Since neither in fact contains LNI, the essentialist ought to insist that they should be augmented so as to contain it; certainly the resulting axiomatic extension would be consistent. Alternatively, if the essentialist insists that LNI is not to be added to either T or S4, then they ought equally to insist that LI be divested of its theoremhood, which would require a modification of one of the identity axioms. But in that case, as we have already seen, the essentialist faces the problem that the Barcan-Kripke proof of LI ought to be acceptable to them by their own lights, and the further problem that rejection of that proof involves rejection of haecceities, which is the very thing essentialists need to stymie Chisholm's argument. Furthermore, in rejecting LI the essentialist endorses the possibility of contingent identity. And this is to endorse the incoherent. So any system of modal logic that rejects LI and LNI ought to be dismissed. Moreover, since any system that **endorses** LI and LNI runs foul of inconsistency, as is shown by Chisholm's argument, no system of modal logic is tenable. Thus the essentialist is left no choice but to meet Chisholm's argument head on.

That is, they must reject either the possibility principle or the transitivity principle. But since, as I shall argue, neither should be faulted by the essentialist given their seminal intuition that things could have been otherwise, Chisholm's argument forces the conclusion that essentialism is incoherent. For in as much as the notion of a thing's possibly being otherwise - or in terms of possible worlds, the notion that a thing exists in multiple worlds - is shown to lead to incoherence, the notion of contingent properties equally leads to incoherence, since this latter notion rests on the former. But then any purported distinction between contingent properties and essential properties evaporates. The distinction is simply without content. This

holds of course providing that one endorses PIOICP, the Principle of the Identity of Indiscernibles in Cases of Counterfactual Possibility.

(Admittedly the essentialist could also take the desperate step of denying the assumption of our reductio that there are ever two entirely contemporaneous individuals a and b. (See p.78 of this thesis.) But not only is the denial of this assumption implausible, I could (causally could) easily refute it (the denial): give me two plastic spoons and I will consign them simultaneously to the flames. In any case, it is enough for my purposes that it is logically possible, even if factually false, that there are two such individuals. The problem for essentialism then becomes that it commits one to the **coherency** of the notion of the one individual's possibly having been the other.)

Before concluding this section there is one consideration against the necessity of distinctness which I want briefly to address and dismiss. This is Prior's thought experiment of the person who reproduces like an amoeba, producing two products a and b by fission, each of whom has a perfect memory of having been the one original person, though now each is being and doing quite different things. (Prior 1957, p.64.) In such circumstances Prior thinks that it would be accurate for a to say to b 'Once you were me' and *vice versa*, although neither could say accurately 'Now you are me'. But then a and b, while distinct **now**, have **not always** been so. Thus a and b are **not necessarily** distinct, for in general one would think that if x and y are necessarily R-related then surely they are R-related at all the times of their respective histories. My response to this argument is to maintain that neither a nor b could say correctly to the other 'Once you were me'. For suppose that both a and b have a memory of eating sardines, say. Then as Prior points out, we can use the escape-hatch of saying that neither a's nor b's memory of the sardine-eating is veridical - the person who ate the sardines was someone other than either a or b,

and is now non-existent. (Prior 1957, p.65.) In short, a and b each came into existence only with the fission and destruction of the parent. So they couldn't have been identical before this fission. Certainly in the case of unconscious organisms such as amoeba there is no inclination to describe the offspring as having been identical before the fission of the parent.

2.3 AN ESSENTIALIST REPLY

2.3.1 REJECT THE IDENTITY OF INDISCERNIBLES FOR COUNTERFACTUAL POSSIBILITIES

The essentialist might try to meet Chisholm's argument through a denial of PIOICP. And yet on one reading of this principle, it would appear to be straightforwardly true. For if a's being identical with a - recall this is a's haecceity - is allowed to count as a genuine property of a, then if an object named 'b' could have had all of the properties of a, then b could have had a's haecceity and thus the property of being identical with a. So if b could have had all of a's properties, b could have been identical with a. Let us call the thesis that an object's haecceity is a genuine and primitive, i.e. non-composite or irreducible, property of that object, 'haecceitism'. The essentialist might attempt to deny haecceitism in order to evade PIOICP; they might reply that a's being identical with a is not a genuine primitive property of a in the way, say, its height is - a reply with which I am certainly sympathetic. But unfortunately for the essentialist I shall now present another argument for PIOICP that does not trade on the genuineness of haecceities. This argument, if successful, will force the essentialist to train their guns on either the possibility principle or the transitivity principle; and I shall argue that the essentialist will then have to **endorse** haecceitism in order to refute either of these principles. In fact, essentialists often endorse haecceitism - though I very much doubt that it is out

of an appreciation of the difficulties I shall presently raise. How then do they evade the argument at the beginning of this paragraph for PIOICP? Simply by denying that b **could** have had a's haecceity, which is thus to say that b could not have had all of a's properties. For further discussion of this point see 2.3.2.1.

The argument for PIOICP is simply an appeal to consider the mysteriousness of its denial. After all, there is something deeply puzzling in the prospect of the universe's possibly having had exactly the genuine properties it has, except for the circumstance that, say, I don't exist, but rather there exists a person exactly like me in every genuine property, who yet is distinct from me. Or again there is something deeply puzzling in the suggestion that our universe may not have existed, but a **distinct** qualitatively identical universe instead. For just what would have made this other universe distinct from our universe? I do not mean to suggest that no two **actual** objects could never be indiscernible, for if neither its haecceity nor its spatial position counts as a property of an object, then I do not see why, say, two actual spatially separated electrons could not have all their properties in common. What I do mean to suggest is that in cases of counterfactual possibility, the suggestion, viz. that a non-actual object qualitatively identical to an object might have existed in its place, seems deeply mystifying. What after all would have made it **that** object and not the actual object?

On this latter point the essentialist Monte Cook concurs substantially. It is worth quoting him at length:

> Here the obvious question is what makes what is seemingly one object different objects. And here, clearly, nothing intrinsic to the objects guarantees that they are different. ... If no intrinsic feature of an object determines whether in different worlds we have that object or another one, then [presumably] some extrinsic feature must do the job. According to some philosophers, what makes a 'temporal slice' a temporal slice of a particular object is not something intrinsic to that slice, but some causal or spatio-temporal relation it bears to other slices that somehow constitute the object. But this sort of relation cannot make different 'world slices' world slices of the same object, because causal and spatio-temporal relations do not hold

between possible worlds. These familiar ways of guaranteeing that we have the same object, then, do not apply to possible worlds.

It may seem that I have just been begging the question. I've assumed that something must make an object in one world the same as or different from an object in another world. But assuming this may seem to be assuming precisely what ... [the opponent of PIOICP] wants explicitly to deny. [They are] ... arguing for the possibility of indiscernibly different objects. All the same, ... [they have] to in some way make plausible the claim that we have different objects; ... [they] can't ... just assert it. Since ... [they] have to appeal to some intrinsic or extrinsic feature of the objects to make ... [their] claim plausible, ... [they] can't say that nothing need make an object in one world the same as or different from an object in another world. ... The problem ... is that [the opponent of PIOICP has] deprived us of any way of making it plausible that individuals in different worlds are or are not the same. There is no story that one can give to back up such claims. The story ... [they] try to give us is in the end really no story at all; it reduces to unsupported stipulations that individuals in different worlds are or are not the same. (Cook 1987, pp.504-506.)

Put it this way. What in the end is the difference between the circumstance in which it is Adam who has all of Noah's properties and Noah who has all of Adam's properties, and the circumstance in which it is Noah who has all of Noah's properties and Adam who has all of Adam's properties? If you say 'there is no difference' then I take this to support PIOICP. While Cook in effect appeals to PIOICP in his advocacy of the essentiality of origin, he fails to see that the same consideration leads to the conclusion that an object could have been different in **any** respect, not just respect of origin, only on pain of absurdity.

Essentialists themselves are divided on the issue of PIOICP. Thus Forbes in effect endorses PIOICP with his repeated insistence, in *The Metaphysics of Modality*, that there can be no what he calls 'bare identities', which is to say that the distinctness of objects must be grounded in **real ,** i.e intrinsic, differences between them. Forbes, unlike Cook, does not countenance the possibility of extrinsic determinations of transworld identity. And even an agnostic about PIOICP such as Stalnaker confesses that there is something both bizarre and implausible about the assumption that an individual might have had all the properties of another. (Stalnaker 1979, p.349). Interestingly, Lewis ought to endorse PIOICP, given his views on the

relation of constitution. (See p.96 of this thesis, n.11.)

Salmon in contrast waives PIOICP at the cost of renouncing the transitivity of the accessibility relation between possible worlds, which is thus to abandon S5 and S4. (I explore the rationale of Salmon's stance in chapter 3.) Salmon replaces PIOICP with a principle to the effect that two supposedly distinct possible worlds are in fact identical just where the very same **facts** obtain in each world. So according to Salmon, two worlds may be materially indiscernible at every moment of their respective histories, yet be distinct worlds because different facts obtain in each; e.g. in one world the fact obtains that object a occupies the space-time region (δx, δy, δz, δt), while in the other world the distinct fact obtains that object b occupies the same space-time region, even though the matter filling the respective regions is identical. (Salmon 1986, p.108.) I find this a dubious suggestion, not least of all because it seems to make the identity of facts ontologically prior to the identity of the objects that the facts are about. Furthermore, I can make little sense of the notion of **conditional** possibilities that is entailed by the rejection of the transitivity of the relation of accessibility between worlds. As Carter observes, conditional possibilities are intelligible with reference to causal possibility, but not to metaphysical possibility. I shall make good this claim in the next chapter, in section 3.7.3.2.

In a more recent paper Salmon has taken a different tack against PIOICP:

> [There is a] philosophically important metaphysical thesis that, where x is a possible individual from a possible world w_1 and y is a possible individual from a possible world w_2, if x=y then there must be something in the qualitative nature of x and y, as they are in these worlds, that makes this so, some fact about the qualitative character of x in w_1 and y in w_2 in virtue of which they are identical. This thesis (which is one of various theses that go by the name 'anti-haecceitism') is false. In fact, despite its popularity and *prima facie* appeal, precisely the opposite is (virtually) provable: where x is a possible individual from a possible world w_1 and y is a possible individual from a possible world w_2, if x=y then there is no fact about their qualitative character (as they are in these worlds) in virtue of which this is so, and there is nothing in the qualitative nature of x and y, other than their mere possible

> existence, that makes them identical. For surely there is no qualitative fact about x, other than the fact of its possible existence, in virtue of which x=x. That is, x is such that there is nothing in its qualitative character (in any possible world) that makes x identical with it. It follows by Leibniz's Law that if x=y, then y is also such that there is nothing in its qualitative character that makes x identical with it. Therefore, if x=y, then there is nothing in x's, i.e. y's, qualitative character that makes x=y. Q.E.D. (Salmon 1987, p.517.)

Now it might be true, as Schlesinger observes, that the reason that an individual is identical with itself is not because it has various qualities but because the very concept of identity demands that everything be identical with itself. (Schlesinger 1994, p.197.) Nonetheless Salmon's argument does not speak to the issue of the **distinctness** of individuals. And it is the distinctness of x from y, not the identity of x with itself, which is the more important issue in assessing PIOICP. Now the reason surely that one object is **distinct** from **another** object is that the one has properties which the other lacks, or the one lacks properties which the other has. In any case, the distinctness of actual objects in actual situations is grounded in the respective qualitative character of the distinct objects in question. Hence it is plausible to assume that the distinctness of possible objects in counterfactual situations is similarly grounded in facts of respective qualitative character. So to suppose that an object a could have had all the properties of a distinct object b, and yet been a and not b, would be to violate the fact that distinctness is grounded in the qualitative character of the individuals said to be distinct.

For her part, Penelope Mackie grants that there

> does seem to be something odd about the idea that there could be two possible worlds such that one world contains, say, one particular rat, whereas the other contains , in its place , a distinct but indiscenible rat, where there is absolutely no difference , in respect of properties, or the identity of matter, or the identities of other individuals, that one could appeal to to back up the claim that these possibilities are genuinely distinct. (Mackie 1987, pp. 181-182).

Mackie concedes that an appeal to Forbes' No Bare Identities principle (my PIOICP) together with the thesis that the identities of things cannot be extrinsically

determined (see chapter 3, section 3.9) suggests that the idea of counterfactual possibilities is incoherent if individuals lack distinctive essential properties that distinguish them from all other individuals. (Mackie 1987, p.182). Nonetheless she thinks that conceding the incoherence of talk of counterfactual possibility is a dismal option (Mackie 1987, p.190), since it is counterintuitive:

> what we are confronted with seems to be a case where we just **cannot** have everything that we want. Some people, refusing to compromise, will see an incoherence in our ideas about *de re* modality that forces us to abandon the whole enterprise. But if, as is surely reasonable, we prefer to avoid this extreme position, the question becomes one of deciding which of the available compromises involves the least sacrifice. I maintain that the best that we can do is to deny the thesis that there can be no bare difference between identities across possible worlds. (Mackie 1987, p.198)

According to Mackie, the only cost of denying PIOICP is a commitment to certain consequences in tension with our initial intuitions about possibility, but then we can't have everything, so she says. (Mackie 1987, p.198) In particular, there is a commitment to the consequence

> that (assuming the characters in Beatrix Potter's story to be real rather than fictional rats) it is coherent to suppose that for every combination of properties that Samuel Whiskers actually has, it is a possibility that some quite different rat (Anna Maria, perhaps) should have had that very same combination of properties [leaving aside dubious properties such as identity with Whiskers or non-identity with Anna Maria.] (Mackie, P. 1987, p.182.)

As one who refuses to compromise with what I take to be unintelligible I plump instead for the complete abandonment of talk about counterfactual possibility. Perhaps the price is not as high as one might think, for it seems to me that what principally motivates the desire to allow the respectability of such talk is the thought that if things could not have been otherwise then we do not have free will. But a causal determinist who is a compatibilist about free will and determinism will not allow that things could have been otherwise in a causal sense of 'could have'. So why insist that things could have been otherwise in a metaphysical sense of 'could

have', if one's main concern is to secure the existence of free will?

There is a further consideration in favour of my dispensing with talk of things' possibly having been otherwise, which might help to break the impasse here. That is that Mackie saves talk of things' possibly having been otherwise at the expense of a rejection of **full-essentialism**: no ordinary properties are essential to any actual object since any actual object could have been anyhow. But from what then does talk of things' having possibly been otherwise derive its theoretical interest if full-essentialism is in fact inapplicable to reality? What is the point of the promiscuity of alternative possibilities if there are no constraints on such possibilities? Stalnaker similarly muses:

> I cannot think of any point in making the counterfactual supposition that Babe Ruth is a billiard ball; there is nothing I can say about him in that imagined state I could not just as well say about billiard balls that are not him, and if he does have the ... potential to be a billiard ball, it is of no interest that he does since on the ...[theory that anything could have been anything] this does not distinguish him from anything else. (Stalnaker 1979, P.350)

Nevertheless this consideration is far from compelling, since as Stalnaker observes:

> that this alleged possibility [concerning Babe Ruth] is uninteresting is not sufficient reason to judge it impossible...since it is hard to separate the intuition that a proposed possible situation is too bizarre to be interesting from the intuition that it is too bizarre to be possible. (Stalnaker 1979, pp.350-351)

I have argued that, in so far as sense can be made of the statement that a could have had all of b's properties, one ought to conclude that a could have been b. It follows that in their quest to disarm Chisholm's argument for the incoherency of the notion of contingent properties and thereby of essentialism, the essentialist should turn their critical attention on either the possibility principle or the transitivity principle, given their failure to dispatch PIOICP satisfactorily.

2.3.2 REJECT THE POSSIBILITY PRINCIPLE

Suppose the essentialist were to reject an unrestricted form of the possibility principle. They would urge that for some, perhaps even many properties F, a could have had all of b's properties if a and b differ only in so far as a has G which is incompatible with F where b has F. But they would insist that there is at least one particular property F, such that even if a and b differ only in that a has G where b has F, still a could not have had all of b's properties, because a could not have had F. So the possibility principle would not hold indiscriminately for any property constituting the sole difference between a and b. On this view, indiscriminately iterated applications of the possibility principle suffer the same fate as other sorites paradoxes: while a heap of sand minus a grain of sand remains a heap, repeated removals of grains of sand will eventually annihilate the heap. Similarly while objects can tolerate some degree of so-called exchange of properties without detriment to their respective identities, they can't tolerate a total exchange of their properties. But to bolster this approach the essentialist will have to supply an account as to which properties violate the possibility principle and why. The account as to why is simple enough: the properties which violate the possibility principle do so because they are essential to an object's identity, i.e. without those properties the object ceases to be. On this view, an object retains its identity under change of certain properties, viz. its contingent ones, but not under change of other properties, viz its essential ones. As for the account as to which properties violate the possibility principle, it must - if it is not to be vacuous - specify the essential properties peculiar to an object and explain why just those properties are essential to it.

But it is just on this score that we shall find the essentialist on shaky ground in the following chapters. For of all the properties we supposedly could have had how do

we **know** which properties are the ones we could have had and which we could have lacked, or which properties are the ones we could have been without and which we could not have been without? As Cartwright observes:

> Chief among ... [the perplexities of essentialism] is the obscurity of the grounds on which ratings of attributes as essential or accidential are to be made. Apparently, in any particular case, one is simply to reflect on the question whether the object in question could or could not have lacked the attribute in question. Now, such reflection perhaps assures us that every object is necessarily self-identical and that 9 is necessarily greater than 7. But the criteria to which one appeals in such reflection are sufficiently obscure to leave me, at least, with an embarrassingly large number of undecided cases. Is Dancer's Image necessarily or contingently a male? Necessarily or contingently a thoroughbred? And, for that matter, necessarily or contingently not a philosopher? The existence of such cases, even in such large number, does not show that there simply is no distinction between essential and accidental attributes of an object. But it does show that the distinction is a good deal less clear than essentialists are wont to suppose. (Cartwright, pp. 158-159.)

In a similar vein, Plantinga asks:

> is there a world w and an object x existing in w such that x is identical with Socrates and x ... was born in 1500 B.C. or was an eighteenth century Irish washerwoman? Is there a world in which there exists something identical with Socrates and the President of the United States? These questions ... are questions as to which of Socrates' properties are essential to him. Could he have had the property of being disembodied -at-some-time-or-another? Or the property of having-an-alligator-body-at-some-time-or-other? ...These are ... questions ... we may be unable to answer ... It would be overzealous, however, to conclude that there is **difficulty or confusion in the idea that Socrates exists in many possible worlds** [my emphasis]. Our lack of confident answers means only that Socrates has **some** properties such that we cannot easily tell whether or not they are essential to him; it does not so mush as suggest that all his properties are thus inscrutable. And, indeed, not al of them are. Obviously Socrates could have been a bit shorter or a bit taller, a bit wiser or a bit less upright. But then Socrates is not a worldbound individual; there are many possible worlds in which he exists (Plantinga 1974, p.101.)

According to Plantinga, Socrates' **scrutably** essential properties include those rather artificial seeming or extraordinary ones such as: self-identity; being coloured if red; being something or other; being either a prime number or something else;

being unmarried if a bachelor; having some properties; and existence (Plantinga 1974, pp.60-61) - properties, which according to the full-essentialist, **every** actual object would have essentially if any did. Furthermore, among Socrates' scrutably essential properties are those which **many,** if not every actual, object would have essentially, properties like e.g. being a non-number or being possibly conscious (Plantinga 1974, p.62). Now the full-essentialist will standardly urge that these sorts of properties are not **uniquely** essential to Socrates. Moreover, they will standardly concur with Plantinga's claim that Socrates' ordinary meat-and-potatoes properties, such as: having been born in 470 B.C.; being an Athenian; being an philosopher, having lived for some seventy years; having been executed on a charge of corrupting the youth; being the teacher of Plato etc.; - the last of which **is** unique to him - are **not** essential to him, either singly or disjunctively (Plantinga 1974, p.61), and *a fortiori* not **uniquely** essential to him. But then what property does there remain to characterise Socrates both essentially and uniquely? For if Socrates has no essential property which characterises him **uniquely**, then Chisholm's argument stands. (cf. Chisholm 1967, p.84) It is not enough for the essentialist that Socrates have some properties essential to him, for they would standardly concede that other individuals could equally have those properties essential to them, so that if a and b were to exchange - so to speak - an essential property they had in common, no harm would be done to their identities as a and b respectively. What is required to stymie Chisholm's argument is that Socrates have some property which is not only essential to him but **unique** to him. In default of such a uniquely essential property, restrictions on the possibility principle would be arbitrary. But given an unrestricted possibility principle in conjunction with PIOCP and the transitivity principle, the notion of an object's existence in multiple possible worlds leads to incoherence.

It follows that Plantinga's observation that Socrates can exist in many possible worlds because he obviously could have been a bit shorter or a bit taller or somewhat wiser or less upright, is beside the point. Unless Plantinga can point to an essential property of Socrates that is **unique** to him, Chisholm's argument againt the coherency of the notion of existence in multiple possible worlds stands. Thus the notion that things could have been otherwise will be fraught with absurdity.

Before we examine the issue of uniquely essential properties we must pause to consider a suggestion made at the beginning of this section that Chisholm's agument is unsound, turning as it does on an iterated appeal to the possibility principle. And this is illicit in as much as it parallels an appeal to similar principles which generate standard sorites paradoxes. Thus for example if a is not bald with 100 000 hairs, then a is not bald with 99 999 hairs; ... ; finally: if a is not bald with one hair then a is not bald with no hair. But since a has 100 000 hairs and is not bald, it follows by an iterated application of *modus ponens* on the appropriate possibility principle that a is not bald when having no hair. The parallel with sorites paradoxes will not hold up however. For in the case of the predicate 'is bald' there are clearcut cases authorised by ordinary usage of what is to count as being bald and what is to count as being hirsute. Trouble for the possibility principle arises at the fuzzy borders of the intermediate zone between baldness and hirsuteness. But in the case of Chisholm's argument, by the mainstream (i.e. non-Leibnizian) full-essentialist's own admission (recall Plantinga above), as far as an individual's **meat and potato** properties are concerned, none of these properties is essential to the individual, either singly or disjunctively. That is to say that Socrates counts as Socrates no matter how few or many of his **ordinary** properties are hypothetically varied. There is thus no distinction as far as **these** properties are concerned between clearcut cases where we have Socrates and where we do not. As far as Socrates' ordinary properties are concerned, all variations count as cases of Socrates. But then there is

no intermediate zone with fuzzy borders between clearcut cases of Socrates and clearcut cases of non-Socrates, paralleling that betweeen 'is bald' and 'is not bald'. If the possibility principle is to be faulted it can be, by the full-essentialist's own admission, only where it makes reference to a uniquely essential and thus unordinary property of the individual in question. It is to the issue of the existence of such properties that we now turn.

(Recall Penelope Mackie's view that the legitimacy of talk of things' having possibly been otherwise does not depend on the availability of uniquely essential properties which individuate the objects that could have been otherwise. Now Mackie concedes that an appeal to Forbes' principle of No Bare Identities in *The Metaphysics of Modality* (my PIOICP), together with the thesis that the identities of things cannot be extrinsically determined (see the next chapter, section 3.9), suggests that the idea of counterfactual possibilities is incoherent if individuals lack distinctive essential properties that mark them off from all other individuals. (Mackie, P. 1987, p.182). Nonetheless, she thinks that conceding the incoherence of talk of counterfactual possibility is a dismal option (Mackie, P. 1987, p.190), since it is counterintuitive. Her response is to deny PIOICP through a denial of the requirement of uniquely essential properties. I suggest that the essentialist would do better to vindicate Mackie's claim of counterintuitiveness in a more plausible manner, namely by grasping the nettle and establishing the existence of uniquely essential properties.)

2.3.2.1 HAECCEITIES AS UNIQUELY ESSENTIAL

Consider the property **being Socrates** or **being identical with Socrates**, and thus the property **being Socrates or Greek**, etc. (Plantinga 1974, p.62). This alleged property of **being (identical with) Socrates** is of course Socrates' haecceity.

Clearly if Socrates' haecceity is a genuine and primitive (irreducible) property of Socrates, it would be not only essential to him but **unique** to him - unlike the property of being self-identical. Thus Socrates could not exchange - so to speak - **all** his properties with Plato and yet survive so to speak as Socrates 'gaining' all of Plato's properties, for in the so-called exchange Socrates would have to lose his haecceity: but then the object that would have all of Plato's properties would not be Socrates. Given the existence of haecceities, the essentialist could foil Chisholm's argument, in as much as Chisholm would not be entitled to his claim that if a and b each exist in more than one possible world, than a and b can exchange - so to speak - **all** their properties. And it is on this claim of course that the difficulties we pointed out depend.

So the question the essentialist must answer is whether haecceities exist. Plantinga's view is that all of Socrates' ordinary meat-and-potato properties are inessential to him both disjunctively and conjunctively, and seems representative of the standard full-essentialist position. So Socrates' haecceity - if it exists - will not be compounded from these garden variety properties. Nor is it clear how Socrates' haecceity - if it exists - will be compounded from those less ordinary properties he shares with other actual objects, such as: being coloured if red; existence; being a non-number etc. So Socrates' haecceity will have to be a primitive or non-composite property. Are there such properties? I have clearly indicated my inclination to doubt this. My doubt is motivated by the thought that haecceities - if they exist - are nothing over and above the conjunction of an object's **ordinary** properties, the very kind of properties that Plantinga would urge are inessential to their bearers, either singly or disjunctively combinatorially. It is these ordinary properties which do all the ontological work of individuation. If my inclination is correct, then Chisholm's objection wins the day, provided of course that we uphold the Principle of the Identity of Indiscernibles in Cases of Counterfactual Possibility.

But to vindicate my position on haecceities would require an examination of the issue of what it is to be a genuine primitive property, an issue beyond the purview of this thesis. Since I am unable to undertake this examination in this thesis, the essentialist can claim a stand-off.

For his part, Adams is one essentialist who would reply that this stand-off can in fact be broken in their favour, for he advances an argument for the existence of primitive haecceities. He wants to say that not only are there such properties as haecceities, but that they are non-derivative: that is to say that they are not constructed out of what he calls 'general properties', i.e. properties that are capable of being possessed by different individuals (Adams 1979, p.7). Now at the very outset Adams overlooks a crucial ambiguity in the phrase 'capable of being possessed by diferent individuals': it is amenable to both a metaphysical interpretation and a logical one. Firstly, it could mean that it is **metaphysically** possible that the property in question is possessed by different individuals, even if in reality it is uniquely instantiated. Secondly, it could mean that there are at least two objects such that for each object it is **logically** possible to assert of it - i.e. there is no self-contradiction in asserting of it - when singled out referentially, that it possesses the property in question. In this second sense the property of being red is general because there are in fact at least two red objects, from which it follows that it is not self-contradictory to assert of each of these objects that it is red. Alternatively, the property of being red is general in this second sense because there are in fact at least two green objects such that when each is picked out referentially - as opposed to attributively, e.g. as that green object - it is not self-contradictory to assert of each that it is red. It is not contradictory but merely false. The property of being the world's tallest building is also general in this second sense, because while it is logically necessary that it is possessed by only one object at any one time, there is no self-contradiction in attributing it to a building that in fact fails to possess it, and since there is no self-

contradiction in attributing it to its actual possessor, it follows that there are at least two objects such that for each there is no self-contradiction in asserting of it that it has the property of being the world's tallest building. Contradiction arises only when both attributions are conjoined. Clearly, an anti-essentialist could grant the coherency of this second sense of 'general properties' even though they would want to reject as incoherent the first metaphysical sense.

Adams wants to contrast general properties with haecceities. But the contrast does not hold for both senses of 'general'. Take for example the property of being identical with Socrates. In the first sense of 'general', Socrates' haecceity is not general for Adams, presumably because he would say that it is metaphysically impossible for any object distinct from Socrates to be identical with Socrates. But in the second sense of 'general', Socrates' haecceity is general, because it is not self-contradictory to assert of something singled out referentially which is not identical with Socrates, that it is identical with Socrates. It is no more self-contradictory than an assertion of the Adelaide Hilton when singled out referentially that it is the world's tallest building. If we dropped the proviso that the object be singled out referentially, then in as much as it is self-contradictory to say 'This building which is not the world's tallest building is the world's tallest building', the property of being the world's tallest building would count as **non-general**. But this would be contrary to the spirit of Adams'distinction between general and non-general properties, for he wants to oppose haecceities to all **general** properties. So to stay faithful to his intentions we shall have to understand 'general properties' in the first sense of 'general', i.e. as properties which are **metaphysically** capable of being possessed by distinct individuals. On this view, an individual's haecceity is metaphysically incapable of being possessed by individuals distinct from it, even in a possible world in which the individual in question does not exist - in contrast to a property like being the world's tallest building; for while this latter property cannot

be instantiated by more than ond individual in any world, it can be instantiated by a different individual from world to world. But this view of haecceities is just what the essentialist needs to falsify an unrestricted possibility principle, and thereby to meet the gravamen of Chisholm's argument that, in as much as the notion of contingent properties is incoherent, essentialism itself is incoherent.

A point of terminology: where I have spoken of general properties Adams also speaks of qualitative properties or suchnesses. The terms are meant to be synonymous. Where I have spoken of haecceities Adams also speaks of thisnesses. We may now present Adams' argument for the existence of primitive haecceities:

> suppose the [principle of the Identity of the **Qualitatively** Indiscernible] ... is true [i.e.no two individuals are exactly alike in all qualitative respects]. And suppose ... that it is true of possible worlds as well as of individuals, so that no two possible worlds are exactly alike in all qualitative respects. Then for each possible individual there will be a suchness of the disjunctive form: having suchnesses Si1 in a world that has suchnesses Sw1, or having suchnesses Si2 in a world that has suchnesses Sw2, or ... which that individual will possess in every world in which it occurs, and which no other individual will possess in any possible world ... . This suchness will, therefore, be necessarily equivalent to the property of being that individual, and, since there will be such a suchness for every individual, it follows that every individual's thisness will be equivalent to a suchness ... . But ... [then] it will be hard to show that thisnesses distinct from suchnesses cannot be dispensed with, or that possible worlds cannot all be constituted purely qualitatively [presumably because of considerations of ontological parsimony].
>
> On the other hand, if it is possible for there to be distinct but qualitatively indiscernible individuals, it is possible for there to be individuals whose thisnesses are both distinct from all suchnesses and necessarily equivalent to no suchness. And in that case there is some point to distinguishing the thisness of individuals systematically from their suchnesses. For it is plausible to suppose that the structure of individuality is sufficiently similar in all cases that, if in some possible cases thisnesses would be distinct from all suchnesses, then thisnesses are universally distinct from suchnesses - even if some thisnesses (including, for all we know, those of all actual individuals) are necessarily equivalent to some suchnesses. (Adams 1979, pp.12-13)

Adams then goes on to argue that the principle of the Identity of the Qualitatively Indiscernible, henceforth IQI, is false: it is not true in general than an object designated 'a' is identical with an object designated 'b' just because a and b share all their qualitative (i.e. general) properties. Adams concludes from this that in general

what individuates an individual is its primitive haecceity.

Adams argues against IQI by means of two thought experiments, the one involving the spatial separation of two objects, the other the temporal separation. Turning to the first: since we can supposedly imagine a universe consisting solely of two distinct objects which are qualitatively identical, what differentiates them must be a non-qualitative, i.e. a non-general, property - in short a thisness or haecceity. Similarly, we can imagine a perfectly cyclical universe in which each event (process or object) is preceded and followed by infinitely many other events (processes or objects) qualitatively identical with it. Again what must be differentiating them is a non-general property, a haecceity.

Now it is plausible enough that what is imaginable is a reliable index of what is logically consistent - at least in this context. Since the situations described are imaginable, their description appears logically consistent. Hence if individuation is a uniform affair - as Adams thinks it reasonable to assume - then if what individuates objects in the situations described are different haecceities, then in general what individuates objects are different haecceities. The question now is whether the situations described have been correctly described as ones in which distinct objects are **qualitatively** identical. Consider for example the fact that the objects are spatially or temporally separated. To be thus separated they must occupy different positions in space or time. But on a non-relational substantivalist view of space and time an object's occupation of a spatial or temporal position counts as a general and thus qualitative property of the object. It is at any rate not clear why it should not so count. In that case the objects described as spatially or temporally separated are incorrectly described as qualitatively identical, for they are not: they have different spatial and temporal psitions. On a relational view of space and time, position in space and time will be defined in terms of certain reference objects or events in

space and time. But again it seems that two spatially or temporally separated objects will differ in being differently related to the reference object. And again it is not clear why such differing relations should not count as **properties** of the objects in question. In any case IQI is standardly understood to encompass not just monadic attributes but properties of higher adicity, i.e. relational properties. Moreover these differing relations will count as **general** or qualitative properties of these objects. So again a description of the two objects as qualitatively identical is incorrect.

I do not claim to have shown conclusively than an object's occupation of a position in space or time counts as a (relational) property of that object. But it does seem to me that Adams just assumes the opposite. If we deny his assumption, and I think that such a denial is plausible, then Adams has provided no reason why we should follow him in his description of the situations under discussion as ones in which two objects are qualitatively identical. Thus I see no reason to concede the existence of primitive haecceities. The stalemate over haecceities remains, and so Chisholm's argument remains unrefuted; the onus remains on the essentialist to vindicate their position.

2.3.2.2 THE ESSENTIALITY OF UNIQUE WORLD-INDEXED PROPERTIES

We now turn to another candidate for the position of being a uniquely essential property of Socrates. Specifically, I have in mind Plantinga's notion of world-indexed properties. According to Plantinga, these are essential to the individuals that bear them. Furthermore, if they are allowed to count as properties, then there would seem to be some which are uniquely instantiated. So the anti-essentialist will have to consider Plantinga's claim that world-indexed properties are essential to

their bearers.

Take a property that Socrates actually and uniquely has, say, the property of being the teacher of Plato. Call it 'F'. Now Plantinga construes possible worlds as maximally consistent states of affairs (Plantinga 1974, p.45), i.e. if w is a possible world, and S is a possible state of affairs, then either w implies S or w implies not-S. Let 'α ' denote the actual world. Since α is a possible world, it will be a maximally consistent state of affairs. Call the state of affairs consisting in Socrates' being the teacher of Plato 'S_p'. Then a entails S_p. Now to say Socrates is F in w is to say that w had obtained, Socrates would have been F. But what does it mean to say that Socrates has the world-indexed F-in-a in w? How is a related to w? According to Plantinga, to say Socrates has F-in-a in w, is to say that if w had obtained, it would have been impossible for a to obtain and S not to (Plantinga 1974, p.62). Let us defer for now consideration of the question whether being F-in-a is a genuine property of Socrates. Supposing it is, is it unique to Socrates? Furthermore, is it essential to him? We can readily show that being F-in-a is indeed unique to Socrates in any world w. For imagine some world w_1, such that some a not identical with Socrates is F-in-a in w_1. Call the state of affairs consisting in a's being F in α 'S_{p1}'. Then if w_1 had obtained, it would have been impossible for a to obtain and S_{p1} not to. But by hypothesis it would be impossible for a to obtain and S_p not to - on an understanding that is of 'a' as referring rigidly to the maximally consistent conjunctive state of affairs consisting of all actual past, present and future states of affairs, viz. S_1 & S_2 & S_3 & ... & S_p & ... . Hence if w_1 had obtained, both Sp and S_{p1} would have obtained. Now possible worlds are consistent states of affairs, i.e. if a entails S_p then it does not entail not-S_p. Since by hypothesis Socrates is **uniquely** F, then for S_p and S_{p1} to be compatible, S_{p1} must concern Socrates, i.e. a must be identical to Socrates. Otherwise the set {a, S_p, S_{p1}} would be inconsistent. Thus being F-in-a is unique to Socrates in any world in which he exists.

But this proof hinges critically on its interpretation of 'a' as referring to S_1 & S_2 & S_3 & ... & S_p & ... . For **it** is only if 'a' so refers that it is impossible for a to obtain and S_p not to. If 'a' is taken to mean 'the actual world, however it is constituted', then in saying that it is impossible for a to obtain and S_p not to, the essentialist is committed to the claim that the world, however it would have actually turned out, would have included S_p, or in other words, that Socrates essentially is a teacher of Plato. But this is a claim from which the essentialist would standardly resile. So given this reading of 'a' as referring to S_1 & S_2 &S_3 & ... &S_p & ..., is being F-in-a **essential** to Socrates? In other words, is it true that Socrates is F-in-a in every world in which he exists? Clearly S_1 & S_2 & S_3 & & S_p & ... entails S_p. So for every world w, if w had obtained, it would have been impossible for a to obtain and S_p not to . *A fortiori*, for every world w in which Socrates exists, if w had obtained it would have been impossible for a to obtain and S_p not to. So Socrates is F-in-a in every world in which he exists. But then being F-in-a in every world in which he exists. But then being F-in-a is essential to Socrates.

Has the essentialist found then an answer to Chisholm's argument? That all depends on the acceptability of Plantinga's reading of 'a' as referring to the maximally consistent conjunctive state of affairs S_1 & S_2 & S_3 & & S_p & ... To pick up a question set aside earlier: is being F-in-a a **genuine** property of Socrates? For on Plantinga's reading of 'a', to say that being F-in-a is a property of Socrates is tantamount to saying that it is a property of Socrates that S_1 & S_2 & S_3 & & S_p & ... entails S_p. But if it is a property of Socrates, how does it determine **him** in the way, say, his height does? And if it is a property of **Socrates**, then it should equally be a property of **Napoleon** that S_1 & S_2 & S_3 & & S_p & ... entails S_p. Similarly, if it is a property of Napoleon that he lost the Battle of Waterloo - call the corresponding state of affairs 'S_w' - then it should equally be a property of **Socrates** that S_1 & S_2 & S_3 & ... & S_p & ... & S_w & ... entails S_w. The consideration that we

would be reluctant to recognise a's entailing S_p as a property of Napoleon, or a's entailing S_w as a property of Socrates, suggests that a's entailing S_p is not a property of **Socrates** (and that a's entailing S_w is not a property of Napoleon). But this line of reasoning is inconclusive, for arguably it **is** a property of the number two in a way it is not of the number five that two times three equals six. Nonetheless, I would urge the implausibility of the view that it **is** a property of Socrates in a way that it is **not** a property of Napoleon that a entails S_p. It is interesting to observe in this connection that elsewhere Plantinga admits as plausible the view that one who asserts that God is such that Adam sinned does not thereby predicate a property of God; for *being such that Adam sinned* doesn't really characterise God: *being such that Adam sinned* is relevant to the way **Adam** is; it has nothing to do with **God's** being the way he is. (Plantinga 1980, p.41) By parity of reasoning Plantinga should grant that it is not a property of **Socrates** that α entails S_p., rather it is a property of α, if of anything.

To summarise: if 'a' in 'F-in-a' denotes S_1 & S_2 & S_3 & ... , then being F-in-a is something that is true of Socrates in every world in which he exists, but then it is not clear that this is a **property** of Socrates or thus an **essential** property of him. If on the other hand 'a' does not denote that conjunctive state of affairs but rather has the force of 'actually', so that 'is F-in-a' amounts to 'is actually F', then while it is a property of Socrates that he is F-in-a, the full-essentialist would hardly want to say that the property of being actually F is **essential** to Socrates.

2.3.3 REJECT THE PRINCIPLE OF THE TRANSITIVITY OF PREDICATIVE POSSIBILITY

The upshot of my discussion of an attempt on the part of the essentialist to disarm Chisholm's argument through a rejection of the possibility principle, is that such an attempt founders because of the failure to produce a convincing case for the

existence of properties which are uniquely essential to their bearers. So perhaps the essentialist would be better advised to reject the principle of the transitivity of predicative possibility, i.e. for all F and G, if a could have been F, and an object that is F could have been G, then a could have been G.

But how would a denial of this principle work? The essentialist would have to say that at best it might be true of some or even most predicates 'F' and 'G' that if a could have been F, and an object that is F could have been G, then a could have been G. But it could not hold true for all predicates 'F' and 'G'. For any object, there would have to be some F and some G such that while the object in question could have been F, and an object that is F could have been G, still the object in question could not have been G. So at this point it becomes pertinent to inquire for each particular object which of its properties are such as to nullify the transitivity principle. But here it appears that the essentialist will have to concede that there are no such properties. After all, if - as the essentialist would standardly grant - Socrates could have been a deaf human being, and a deaf human being could have been hunchbacked, it would then seem *ad hoc* for the essentialist to disallow Socrates' possibly having been hunchbacked. But then the essentialist will have to concede that Socrates could have been both deaf and hunchbacked. This line of reasoning eventually leads to the result that Socrates could have had all of Quasimodo's properties, and thus, as I have argued, we ought to conclude that Socrates could have been Quasimodo (assuming for expository convenience that Quasimodo really existed as a contemporary of Socrates.)

Hence there is no good reason why the essentialist should reject the transitivity principle given their seminal intuition that things could have been otherwise. Indeed this intuition commits them to it. Admittedly on a counterpart-theoretic construal of the transitivity principle it comes out false, for counterparthood is based on the

relation of resemblance which is notoriously not transitive. But if counterpart-theoretic formulations are supposed to capture the import of modal statements formulated in the **pre-logical idiom** of ordinary language, the language namely of 'could have' rather than of 'possible worlds', then in so far as it is the **idiomatic** formulations which are the sources of incoherence, I do not see how it is to the point that unidiomatic formulations are appealed to. Thus, as Brody explains, according to counterpart theory,

> an object has a property essentially just in case it has it in the actual world and its counterparts in all possible worlds in which it has one also have that propety. The counterpart in a possible world of an object in the actual world resembles the actual object, and nothing else in the possible world resembles it more...as Lewis himself notes, resemblance in certain respects is obviously going to count more than resemblance in other respects. But what makes one resemblance count more than another? Surely, it is going to be that certain respects are tied up with the object's very nature. But then it is going to be the very same grounds which would determine what are the essential properties of the object. So we get nowhere in trying to understand the essential-accidental distinction by appealing to the idea of a counterpart. (Brody 1973, pp.358-359.)

Hence if the essential-accidental distinction is incoherent, no appeal to counterpart theory will remove the incoherence. Thus an essentialist appeal to the non-transitivity of counterparthood will not help them in their denial of the principle of the transitivity of possibility. This in part is why I reject Forbes' approach to the transitivity problem, which he treats in the narrower context of the essentiality of origin. Forbes (1983) appeals to a semantic apparatus of degrees of truth. But I do not believe that truth comes in degrees. In any case, identity certainly does not, *contra* advocates of vague identity. But similarity does come in degrees, so as Carter suggests, 'the appeal to degrees of truth in response to the slippery slope problem leads directly to, and so suffers no escape from, counterpart theoretic modal semantics.' Counterparts are metaphysicaly and semantically *entia non grata*. (Carter 1986, p.70.) (Counterpart theory has other technical problems; for one, it faces the difficulty of giving an account of verisimilitude which presumably is

required to explicate its notion of degrees of overall resemblance between worlds; certainly the Popperian account of verisimilitude has been criticised by Miller and Tichy. See Mortensen 1983.)

Nor is it to the point to appeal to stipulative concepts such as a ship*, which according to Sainsbury is something which could have been made out of a set of parts differing in at most one member from the set it was actually made of. (Sainsbury 1991, p.277.) If s is a ship* then it could have differed by one part. Furthermore a ship* s_2 , which differed thus from s could have differed from itself by one part. But s could not have had the properties of the resulting ship* s_3 , for it would have had to differ by two and hence more than one part from itself. (Sainsbury 1991, p.277.) But what Sainsbury's example involves is a breakdown of the transitivity principle when applied to **conceptual** possibility. Conceptual possibility I would urge is just a matter of conventionally labelling some things one way and other things another. As such it has no bearing on **my** quarry in this section, viz. the transitivity of **metaphysical** possibility.

Since the three principles underpinning Chisholm's argument, namely PIOICP, the possibility principle and the transitivity principle, should all appear acceptable enough to the essentialist, and since the essentialist's case for the existence of uniquely essential properties is far from compelling, I conclude that Chisholm's argument constitutes a *reductio* of essentialism: as an argument against essentialism it is forceful even if inconclusive.

2.3.4 ARE KRIPKEAN STIPULATIONS A SOLUTION?

Before I close this discussion of the problem for essentialism of a thing's existence in multiple worlds, I must briefly comment on the so-called problem for essentialism of transworld identification. The objection runs that if talk of a thing's possibly

having been otherwise is to make sense, then we must have truth-conditions for statements of the form that a could have been F. But the objection is that we lack such truth-conditions, for we lack a necessary condition of there being such truth conditions, namely that we have criteria for the identification of a from world to world. In the absence of such criteria talk of a's possibly having been otherwise is meaningless. Thus Quine contrasts criteria for transtemporal identification with those for transworld identification:

> our cross-moment identification of bodies turned on continuity of displacement, distortion, and chemical change. These considerations cannot be extended across worlds because you can change anything to anything by easy stages through some connecting series of possible worlds. The devastating difference is that the series of momentary cross-sections of our real world is uniquely imposed on us, ... whereas all manner of paths of continuous gradation from one possible world to another are free for the thinking up. (Quine WA, p.127.)

The essentialist may have recourse to a Kripkean solution, replying as follows:

> the proponent of transworld individuals argues that the critic has been misled by a picture. He supposes that the various possible worlds are spread out before us in all their infinite detail and that before we can talk about a specific individual in one of those worlds, we have to determine, first, whether the individual really exists in the world, and, second, just which individual among the objects existing in the world is the one we wish to talk about. But according to the defender of transworld individuals, this account is misleading ... it assumes that our epistemic access to different possible worlds involves their being given to us for some kind of investigation, as things we can somehow look into ... . [But] our access to possible worlds is ... stipulative; we stipulate just which world is to function as the subject of our counterfactual discourse; ... our stipulation settles questions of identity. We say ... 'Suppose Plato had lived ten years longer'. In saying this, we fasten on a world; but our very way of characterising the world ensures that it is one in which Plato exists and that it is he, among the objects in that world, that we are talking about. Transworld identity ... presents a problem only if we think of worlds as things given us with the identities of their inhabitants as yet unsettled; worlds aren't given us; we stipulate just which world we mean to talk about, and our stipulation ensures that the individual we are concerned with is the object of our reference. (Loux PA, p.44)

The interest for our purposes in a Kripkean stipulative solution derives from the question whether it can deal with Chisholm's argument. I think it cannot. For recall that the problem for essentialism of existence in multiple worlds arises through an unrestricted possibility principle. But it is not clear how Kripkean stipulations will overcome this problem. Firstly, it would be arbitrary to **stipulate** that this uniquely held property rather than that uniquely held property is essential to its bearer; secondly, one could always **stipulate** that despite a complete exchange of properties, a would have remained a and b would have remained b, even though a would have looked exactly as b does and *vice versa*; still the question invites itself: in virtue of what would a still have been a and b still have been b? In default of a satisfactory answer to this question, it seems that the essentialist is just using verbal sleight of hand to meet Chisholm's objection.

2.4 LOOKING AHEAD TO CHAPTER 3

Now, in so far as my argument for the Principle of the Identity of Indiscernibles in Cases of Counterfactual Possibility may not strike the essentialist or the anti-essentialist for that matter as sufficiently cogent, it will be necessary for the anti-essentialist to meet the essentialist half-way. That is to stay, if the anti-essentialist grants the essentialist that Chisholm's argument can be plausibly stymied through a rejection of PIOICP, then they are forced to concede that the distinction between essential and accidental properties may be valid after all. But then the anti-essentialist might want to fall back onto anti-**full**-essentialism. If so, they will have to consider the arguments standardly advanced by the full-essentialist for the more important kinds of allegedly essential properties. Specifically I have in mind the view that an object's origin is essential to its identity, i.e. that any object lacking that origin would not be the object in question; and the view that any object essentially falls under a particular sortal or essentially belongs to a particular natural kind.

Since these are the most important kinds of essentialism standardly espoused by the full-essentialist, it will be incumbent on the anti-full-essentialist unimpressed by my considerations in favour of PIOICP to examine the full-essentialist's arguments for these kinds of essentialism. And a judgement of **these** arguments as flawed will seriously undermine the case for an **interesting** full-essentialism, in as much as the only essentail properties remaining unaffected will be those artificial seeming and ontologically uninteresting ones of being coloured if red; being either a prime number or something else; being at a time t_2 such that one was F at an earlier time t_1, given that one was F at t_1; etc. But such a judgement will of course leave the question of the viability of the distinction between essential and accidental properties untouched. This distinction **will** be threatened, however, if we do insist on PIOICP, for then Chisholm's argument returns with a vengeance. To meet it, the essentialist will have to resort to haecceities, on the issue of whose existence I conceded a stalemate *vis-a-vis* the essentialist. It is to the question of the essentiality of origin that I turn in the next chapter, while chapter 4 will be occupied with the issue of the essentiality of sortal sastisfaction or natural kind membership.

CHAPTER 3 THE ESSENTIALITY OF ORIGIN

3.1	PRELIMINARIES	131
3.2	A COULD BE F; A COULD HAVE BEEN F; A COULD HAVE BECOME F	133
3.3	J.L.MACKIE ON ORIGIN	139
3.4	THE ARGUMENT FROM THE BETTER CANDIDATE	143
3.5	REFERENCE TO FUTURE INDIVIDUALS	144
3.6	MCGINN ON THE ORIGIN OF ORGANISMS	147
3.6.1	IDENTITY OF ZYGOTE WITH ADULT	147
3.6.1.1	ZYGOTE AS DISTINCT FROM ITS CONSTITUTIVE CELL	152
3.6.1.2	THE POSSIBILITY OF SLIGHT DIFFERENCE IS INCOMPATIBLE WITH THE IMPOSSIBILITY OF COMPLETE DIFFERENCE	154
3.6.1.2.1	THE CASE OF MATERIAL COMPOSITION	154
3.6.1.2.2	THE CASE OF GENETIC STRUCTURE	157
3.6.2	GAMETES AND THEIR PARENTS	158
3.6.3	THE ESSENTIALITY OF GAMETES	159
3.7	KRIPKEAN ARGUMENTS FOR THE ESSENTIALITY OF ORIGINAL MATERIAL COMPOSITION	160
3.7.1	KRIPKE'S ARGUMENT	160
3.7.2	SALMON'S ARGUMENT 1	161
3.7.3	SALMON'S ARGUMENT 2	162
3.7.3.1	A PROBLEM FOR V**	166
3.7.3.2	A PROBLEM FOR NOONAN'S PROPOSAL	169
3.7.3.3	EVALUATING NOONAN'S PROPOSAL	172
3.8	AN ARGUMENT AGAINST THE ESSENTIALITY OF ORIGIN	173
3.9	FORBES ON THE ESSENTIALITY OF ORIGIN	176

CHAPTER 3: THE ESSENTIALITY OF ORIGIN

3.1 PRELIMINARIES

Let us suppose then, for the sake of argument, that the distinction between essential and accidental properties is coherent after all; let us suppose therefore that things could have been otherwise, which the distinction presupposes. My concern in this chapter will be with the claim made by some essentialists, notably Kripke, that an object's origin is essential to its identity.

By 'object' I mean a concrete (i.e. spatio-temporal) individual, which may be animate or inanimate, natural or artifactual. (These two dichotomies may cut across each other: thus for example a laboratory - synthesised cell would be both animate and artifactual.) Furthermore, my concern is with concrete individuals understood as **continuants**, things that exist at any time in their entirety and not merely partially, and which exist at **each** moment **over** a period of time, undergoing various changes at various times which they survive as the very same individual, until finally they cease to exist altogether. The reason why I neglect the question of the alleged essentiality of origin in terms of **four-dimensional space-time worms** is that such an approach is commonly thought by **essentialists** to lead to an unduly restictive form of full-essentialism, for it would seem that the shape and location of a space-time worm in four-dimensional space is essential to it. It is then often urged that modal discourse about space-time worms must be conducted in terms of counterpart theory. (See e.g. van Inwagen 1990, pp.252ff.) I repeat my previous misgivings about such a manoeuvre: if counterpart theory is to fulfil its function of capturing the import of ordinary language discourse about possibility and necessity, then it must fail where ordinary language fails. (See p.125 of this thesis.) For his part, Van Inwagen is an essentialist who prefers to reject counterpart theory, which leads him to adopt a 3-D idiom in order to express claims of metaphysical possibility. For my

part, I do not think that four-dimensionalism should be rejected on such a tenuous basis. I would urge that four-dimensionalism has greater plausibility than the notion of metaphysical possibility, so that if the two can be reconciled only on pain of adopting counterpart theory, then so much the worse for both counterpart theory and for metaphysical possibility. In any case, I can see litle plausibility in the notion of alternative metaphysical possibilities, given a four-dimensional view of reality. This is not to endorse however the superiority of four-dimensionalism over three-dimensionalism.

By 'origin' I shall understand the **object** or objects from which the individual in question can be said intuitively to **come**. An organism which is an **animate** individual (whether natural or artifactual) typically has antecedent organisms as its origin: e.g. a pair of amoebae originate by fission from a parent amoeba; a human zygote orginates by the fusion of a male and female gamete; a plant originates from a fertilized ovule. An **inanimate** individual (whether natural or artifactual) typically originates from either animate or inanimate origins: a piece of charcoal (which is natural) comes from a tree (which is animate); a lump of scoria (natural) from lava (inanimate); a table (artifactual) comes from an oak (animate); a glass window (artifactual) from various silicates (inanimate). The thesis of the essentiality of origin is the claim that a concrete individual has essentially (necessarily) the very origin it in fact has. It could not have originated from anything other than its actual origin. If the actual origin had never existed, the individual in question could never have existed either.

3.2. A COULD BE F; A COULD HAVE BECOME F; A COULD HAVE BEEN F

Recall from section 1.3.1 that the metaphysical essentialist would appear to require that the category of the metaphysically necessary truths properly includes the class of logically necessary truths. On the other hand, the metaphysical essentialist would want to say that the metaphysically possible is properly included by the logically possible and that the **causally** or **physically** possible is properly included by the metaphysically possible. By contraposition, the metaphysically **impossible** would be properly included by the physically **impossible.** Now the state of an object at t_1 sets constraints on which changes it is **physically** possible for it to undergo at later t_2. Thus a one-legged athlete will be physically incapable of outrunning a two-legged competitor. Are there similar constraints which the state of an object at t_1 places on the changes which it is **metaphysically** possible for an object to undergo at later t_2?

The essentialist would typically urge that the range of constraints would appear to be considerably narrower than in the case of physical possibility. Thus for example while it is physically impossible for a one-legged athlete to outrun a two-legged competitor, the essentialist would probably allow that it's metaphysically possible in as much as the laws of physiology and anatomy could be different. Similarly, they would say that it appears metaphysically possible for an organism such as a human male to conceive and gestate even though lacking the requisite anatomy (perhaps it spontaneously synthesises an embryo which then develops in the abdomen).

Now my concern just now is with the import of the claims that it is **metaphysically** possible for x to outrun y or for x to fall pregnant. Specifically, are we to understand the alleged metaphysical possibilities as still being in the future at the time of their allegation, or as already in the past? I contend that such possibilities must be understood to be in the **past**. For consider for example the allegation that x

metaphysically could fall pregnant at a time **future** to the allegation. By hypothesis it is not being claimed that for all we know such an event will happen, for that would just mean that the event's occurrence is merely **epistemically** possible, nor is it being claimed that the occurrence of this event is compatible with the conjunction of the laws of physics - provided of course that they remain stable - and the present initial conditions, for that would just mean that the event is **merely causally** possible.

Perhaps then it is being claimed **that the event's occurrence is compatible with possible future changes in the laws of nature** (call this claim 'p'). But what then is the import of 'possible' in the phrase 'possible future changes'? If the essentialist just means that it is **epistemically** possible that the current laws of nature will change, then p is tantamount to the claim that the event's occurrence is compatible with certain laws of nature which while not actually obtaining will obtain for all we know. This then brings us back to an **epistemic** interpretation of p. Suppose the import of 'possible' is that of '**causally** possible'. On this second reading p is tantamount to the claim that the event's occurrence is compatible with certain changes in the laws of nature, which changes are **causally** possible. Now for all we know, the laws of nature have changed and will change again. Even so, it is not clear that it is appropriate to talk of a change in these laws as something that is **causally** possible, for in which sense is it correct to say that it is **compatible** with certain laws of nature that some laws be replaced with other laws of nature? Perhaps it is compatible with some laws that others be changed, but only if they are independent of each other. But the notion of metaphysical possibility is meant to apply equally to laws which are not independent of each other. So if the clause 'which changes are causally possible' is removed from the second reading, we in effect are left with the claim that the event's occurrence is compatible with certain non-actual laws of nature. But then this is just a claim of **logical** possibility, since it

is about which statement is consistent with which. If on the other hand we let stand the troublesome clause 'which changes are causally possible' on the ground that it is admissible to speak of its being causally possible that one set of laws is replaced by another, then p in effect reduces to a claim of **causal** possibility.

Suppose then that we understand the claim that a could become F at some future time t as the claim that a's becoming F is compatible with certain **metaphysically** possible future changes in the laws of nature. But it is this very notion of **metaphysically** - as opposed to merely causally - possible **future** changes that eludes me. Given the analogy of a replay, I can make some preanalytic sense of the notion of the metaphysical possibility of something's **having** happened. But I confess to lacking any analogy that will similarly make intelligible the notion of the metaphysical possibility of something's future happening. Clearly the notion of a replay will not serve this purpose, for in the future case nothing has as yet happened to be replayed with a metaphysical variation so to speak. Since I do not understand any other way in which to explicate the claim that it is metaphysically possible that some **future** event will occur, I conclude that the essentialist ought to restict themselves to past events when alleging the **metaphysical** possibility of some event. What they in effect mean, when they assert the metaphysical possibility of the occurrence of some event e at some **future** time t, is that **after** t they will be able to say that e could have happened. Of course I have already expressed my misgivings about the notion of a thing's possibly having happened in section 1.2.1. Since however my strategy in this chapter is to grant this notion I set aside my misgivings.

To return then to the question which prompted this discussion: does the state of an object at t_1 set constraints on the changes that it it is metaphysically possible for the object to undergo at t_2, in the way that it sets constraints on the changes that it is causally possible for it to undergo? In the light of the preceding discussion this

question must be understood to be whether the previous state of an object set limits on the changes the object could have but didn't undergo at a later time now in the past.

The essentialist would probably answer that if x was F at t_1 then it was metaphysically impossible for x to undergo a change from having been being F to being such that it was not F at t_1. But could x have undergone a change from being F at t_1 to being not-F at t_2? Could it have been the case, for example, that a lead pipe at the bottom of a pond suddenly float to the surface? Again the essentialist would probably answer that it could have in so far as the laws of bouyancy could have suddenly changed. Similarly a one-legged athlete could have outrun a two-legged competitor, and a male could have fallen pregnant, in as much as the laws of nature could have changed to allow for this. Could then the laws of nature have been such as to change so as to allow for the metamorphosis of a frog into a prince?. If the essentialist could grant the preceding possibilities I see no reason why they couldn't allow this possibility as well.

What many essentialists would not grant however is a description of the circumstances canvassed as one in which the very same entity survives the metamorphosis. Wiggins is representative of this view: he would deny the metaphysical possibility of a frog's having become a prince, for 'whether a thing can submit to a certain kind of change or not depends on what it essentially is' (Wiggins 1980, p.216). Wiggins follows Aristotle in his view that organisms essentially (necessarily) belong to natural kinds. Corresponding to these kinds are what Wiggins calls - following Strawson's popularisation of Locke's terminology - 'sortals', general terms associated with principles for individuating at a time an individual to which the sortal applies, and for identifying that individual as the same individual at another (earlier or later) time. On this view, if 'F' represents a sortal,

then whatever is an F is an F at every moment of its existence. Moreover, it is F essentially: the career of an F **must** both metaphysically and causally conform to the sortal's associated principles of individuation and identification. Thus an F metaphysically and causally could not violate these conditions without thereby ceasing to exist. For example, since *human being* is a sortal, if I am now essentially human, there is no time in my **past** at which I **metaphysically** could have been non-human. Similarly, there is no **future** time at which I **causally** could become non-human and survive the change. This is so, according to the sortal essentialist, notwithstanding the fact that it is logically possible because not self-contradictory to assert that I was or am or shall be non-human. Admittedly, the sortal essentialist might grant, it is metaphysically possible for radiation to have transmuted a frog's zygote into a genetically humanoid zygote. Similarly, it might even be causally possible for radiation to transmute the zygote at some future time. But even so, on their view the human's existence would have begun (or will begin) and the frog's ended (or will end) only once the transmutation would have been complete (will have been complete). The developed human could not say retrospectively that they had once been non-human: the non-human zygote did (does) not **become** human; rather, it ceased (ceases) to exist, whereupon a human zygote came (comes) into existence.

How does this view of sortally determined essences, which we shall examine in chapter 4, bear on the question of the essentiality of origin, the thesis that the origin of an individual is essential to it? For his part Wiggins confesses militant agnosticism on this question (Wiggins 1980, p.213). Indeed it is far from obvious that Wiggins' view of sortally determined essences should commit him to the thesis of the essentiality of origin. For even if we grant that an object x which is sortally F can undergo only some changes and not others once **in** existence, why should this restrict the range of possible origins from which x could have come **into** existence,

to just its actual origin? Suppose for example that a radioactively bombarded frog's zygote could have transmuted into a genetically humanoid zygote; the sortal essentialist would probably designate the humanoid zygote in its first moment of existence as the **origin** of the adult human that eventually would have developed from it. To this extent it would appear that on the sortal essentialist view only an object that is sortally F could qualify as an origin for an object which is sortally F. But even so, could not the human which evolved from the humanoid zygote have evolved from a different zygote altogether, e.g. from a zygote that also arose by radioactively induced transmutation of a frog's zygote, but a frog's zygote distinct from the original one? Wouldn't it have been enough that the genetic information was the same, even while the molecules encoding this information were not? If so, then the view of the sortal determination of essence would be consistent with the view that its origin is not essential to a concrete individual. Rather it is the sortally determined **kind** to which an individual belongs that is essentail to the **origin** of the individual. If on the other hand we reject the thesis of the sortal determination of essence - which thesis we shall examine in chapter 4 - then we could reject both the claim that the kind to which an individual belongs is essential to it, and the claim that this kind is also essential to the individual's origin. It would follow that the prince mentioned above could correctly say both that he had once been a frog **and** that nonetheless he could have originated from a non-frog's origin.

Following Kripke, Wiggins points out that the question whether x could have become G given it already exists and is not G, differs from the question whether x could have been G in the sense of having come into existence as a G. (Wiggins 1980, p.215). Given that Nixon, say, was human, could he have been a frog? Could he have come into existence as a frog? On Wiggins' view, once Nixon is (sortally) human, then he must continue to be (sortally) human as long as he exists . His humanity is essential to him throughout his existence. So to make sense of the

question whether Nixon could have been a frog we would have to look to a time t before Nixon came into existence and ask whether it was possible at t that the amphibious organism that would henceforth develop from the fertilised frog spawn to hand would be Nixon. Wiggins would probably reply that an affirmative answer to such a question would render difficult a denial that anything can be anything; and since for Wiggins an acceptance of the thesis that anything can be anything comes at a high conceptual cost (Wiggins 1980, p.214), it would be best to disallow the possibility of Nixon's having been a frog. We shall postpone consideration of this issue of conceptual cost until the next chapter. In any case it is clear that the possibility of Nixon's having been a frog would render inessential to him the actual origin from which he arose (viz. his parents' gametes).

3.3 J.L MACKIE ON ORIGIN

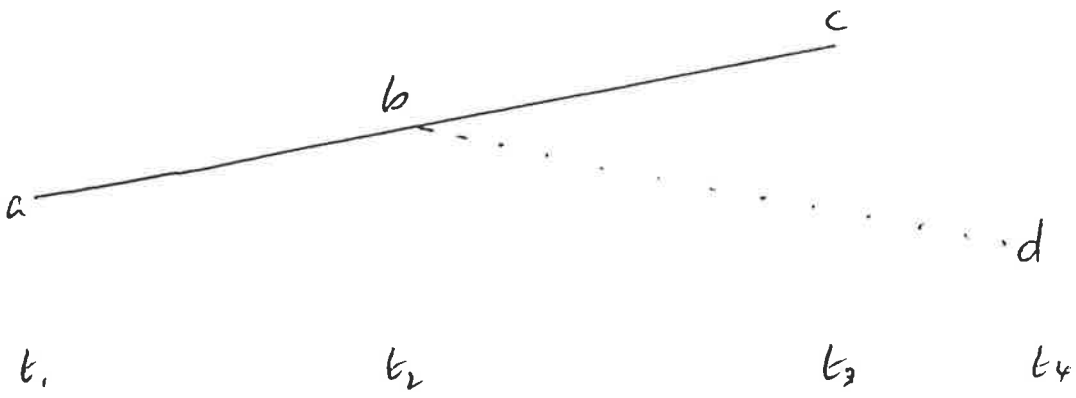
Putting to the side Wiggins' view of the sortal determination of essence, let us consider an argument by J.L.Mackie in Mackie 1974 'De What Re Is De Re Modality?' that an individual's origin is essential to its identity. Unfortunately for the full-essentialist it turns out that despite promising much Mackie in fact delivers very little. His argument is as follows:

> In our ordinary thinking we suppose that, if the future is causally underdetermined, more than one possibility is really open. ... [Now with regard to human action we] take it that there are genuinely open alternatives for the future, but not for the past [as the past is closed. It] ... is fixed just by having occurred (p.558).
>
> [A] ... present state could have come from more than one antecedent [state]...[but] the only way in which it really could have come from some alternative antecedent is by that alternative's being a possible result of some earlier possible divergence from actuality (p.557). [Principle P]
>
> Thus the libertarian view of causal possibilities, conjoined with the view of the past as fixed, ... yield[s the necessity or origin] (pp.558-559).

Now how exactly does the openness of the future conjoined with the fixity of the past yield the essentiality of origin? It is clear to the reader of Mackie that he allows

the following possibility for Nixon, say:

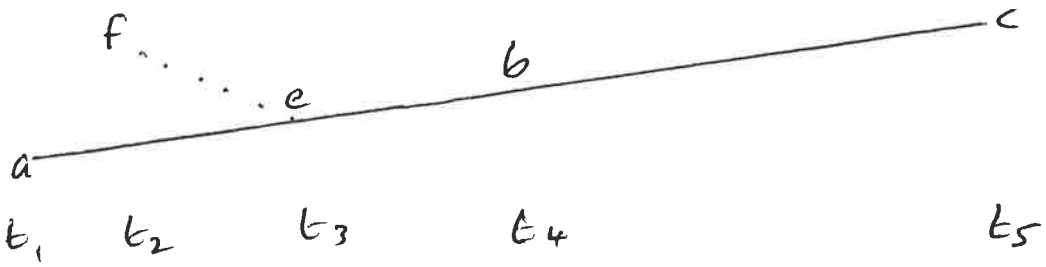
fig. 1



where 'abc' represents Nixon's actual career; 'a' his actual origin; 'b' his actual graduation from university; 'c' his actual end; 'bd' a possible divergence of Nixon's career from actuality; 'd' a possible non-actual alternative end; and 't_1', 't_2', 't_3' and 't_4' the respective times of these various states of affairs.

It is equally clear to the reader of Mackie that he disallows the convergence **fe** in figure 2 as a possibility for Nixon:

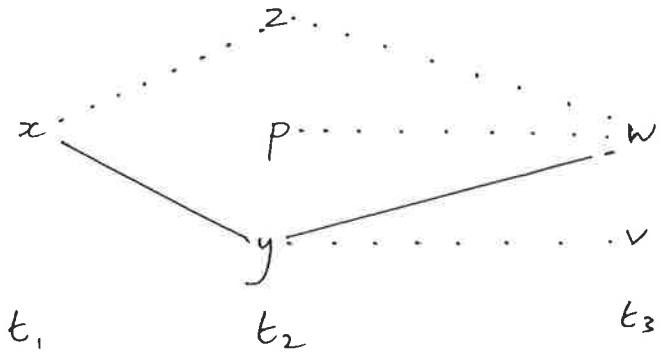
fig. 2



where 'a', 'b', 'c' are as in figure 1; 'e' represents the actual Nixon at the age of 2; 'f' an origin distinct from a; and 't_1' ... 't_5' the respective times.

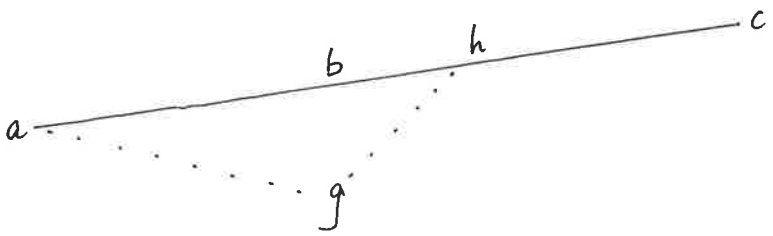
On Mackie's view careers may **diverge** from left to right but not **converge** from left to right except when the convergence is itself a possible result of some earlier possible divergence from actuality. Thus in figure 3

fig. 3



where Nixon's actual career is depicted by 'xyw', 'xzw' represents a possible alternative career for Nixon, since the convergence zw is itself a possible result of the possible divergence xz from xy. The branch 'pw' on the other hand is excluded as a possible alternative career for Nixon, for it is not a **possible** divergence from Nixon's actual career. Mackie must allow for at least some convergences, for if instead he were to disallow as a possible alternative past for Nixon **every** convergence on his actual career path, this would quickly lead to what the essentialist would typically deem to be unpalatable consequences. For typically the essentialist will allow divergences as one proceeds to the right, so that some properties appear on some forks but not on others, thus counting as non-essential or accidental properties of Nixon. In figure 4, for example, Nixon went to university at (b) but presumably need not have: he might instead have got a job in a factory (g).

fig. 4



Nonetheless he also might have quickly become restless and decided to go into politics (h). So if gh were disallowed merely because it is a convergence onto ac, then Nixon **had** to go to university after all: far from being accidental, it was essential to him. But this would strike the majority of essentialists as implausible. This argument can be generalised for any property the essentialist is willing to recognise as accidental to Nixon. So Mackie qua typical essentialist cannot disallow **every** convergence, he must allow some; his solution in fact is to allow just those convergences which themselves

are possible divergences from actuality, which is to invoke principle P.

In this light the impression of *ad hoc*ness that arises on one's first encounter with principle P diminishes somewhat. Principle P then becomes plausible when understood in terms of **physical** or **causal** possibility. For even is a state x earlier then y is causally sufficient for y, it is not the case that y physically could have developed from x if x itself had not been physically possible relative to actual states preceding it. To illustrate: suppose Roosevelt took time off from public life from 1934 to 1936 to train for the Berlin Olympics. Having beaten Jesse Owens in the 100 metres sprint, he returned to Washington to address the problem of the rising tide of fascism (Hitler had snubbed not only Owens). Now while Roosevelt's victory over Owens would have been causally sufficient for Roosevelt's receiving the gold medal, it was physically impossible for Roosevelt to receive this award given his actual physical condition in 1934. So the line of convergence 1934 - 36 onto Roosevelt's actual career would have been physically impossible.

But why should principle P exclude an alternative origin in general as a legitimate possible alternative past for an individual? It is far from clear that every alternative origin would be **physically** impossible relative to the actual states preceding it. Mackie suggests that the answer lies in the fixity of the past. But to say that the past is fixed or closed is to say that Nixon cannot **now** have an origin other than the one he actually has. But this seems to lend no support to the thesis that Nixon could physically not have had a different origin. What could have been but wasn't cannot now have been; nonetheless it might have been, so the essentialist will urge. In the end, Mackie simply seems to be assuming that an alternative origin is physically impossible. And even if it is, would it have been **metaphysically** impossible as well? After all, one who endorses essentialism but opposes **origin** essentialism could suggest that we can imagine Nixon's having developed from a

dried raisin, which upon maternal ingestion bypassed the stomach, lodged itself in the womb, became animate and started to assume a genetic structure identical to that of Nixon's actual zygote. But the opponent of origin essentialism need not appeal to such a *recherche'* possibility in order to meet Mackie's argument. For surely they could simply say that Mackie has failed to provide any non-question-begging explanation for why alternative origins are **physically** impossible. *Ipso facto* he has failed to provide any reason to believe that alternative origins are **metaphysically** impossible. For physical possibility is stronger than metaphysical possibility: event e may be metaphysically possible but not physically so, but the converse does not hold, for if e is physically possible then it is metaphysically possible. So if Mackie's premisses are consistent with the physical possibility of alternative origins, they are *ipso facto* consistent with the **metaphysical** possibility of alternative origins.

3.4 THE ARGUMENT FROM THE BETTER CANDIDATE

At the end of his 'De What Re Is De Re Modality?' Mackie considers an argument against convergence. However the argument can readily be interpreted as an argument against the essentiality of origin and is as follows:

> If we consider a possible world in which some person is conceived at t_0 ..., has a career different from that of the actual Nixon up to t_2 but thereafter just like that of the actual Nixon, we must also be able to consider a possible world in which, as well as **this** person, there is one who is conceived at t_1 [the actual time of Nixon's coneption] by the union of the sperm and ovum that in the actual world produced Nixon, who has the actual Nixon's career for a short time afterwards, but then is ... dropped on his head in infancy and survives only as a human vegetable ... [The latter] has a better claim to be Nixon then ... [the former]. Since there **could** be this stronger claimant, we must not say even with regard to a world from which he is absent, in which we have only the ... [man conceived at t_0], that ... [this man conceived at t_0] is Nixon. (Mackie, J. pp.559-560.)

In short, wherever there are competing claimants for Nixonhood we should always plump for the claimant with the actual Nixon's actual origin; indeed Mackie would

urge that we always do so, competing claimants or no. This argument turns in part on the assumption of the necessity of distinctness, which assumption I argued in 2.2.2 ought to be embraced by the essentialist. Given this assumption, then if b in one world has a better claim to be Nixon than does a in that world, then a is distinct from Nixon in all worlds, i.e. a could not have been Nixon. But the crucial question is why actual origin should be decisive in determining the better claimant. Why for example shouldn't time and place of death be equally decisive? To reply that intuitively one could have lived longer or shorter than one actually did, is to invite the retort that it is not intuitively obvious - at least not to all - that one could not have had an alternative origin to that one actually had.

3.5 REFERENCE TO FUTURE INDIVIDUALS

McGinn observes that an origin - essentialist might want to invoke a principle (call it 'principle O') to the effect that x is essentially F at t_1 if there is no earlier time t at which x exists such that it was physically possible at t that x be not-F at t_1 (McGinn, p.130). Suppose x comes into existence at t_1 in a certain manner. Was there an earlier time at which x existed? Obviously not. *A fortiori* there was no earlier time t at which x existed such that it was physically possible at t that x come into existence at t_1 in a manner other than that in which it did. So by principle O x's actual origin would be essential to x at t_1. McGinn thinks that principle O must be rejected because it has modally extravagant consequences:

> it entails that **everything** true of Nixon [say] at the moment of his creation is necessarily true of him. Not just exact time and place of birth, but also that he started to be in a room containing a vase of geraniums. (McGinn, p.130).

Principle O could also be dismissed by the essentialist on the grounds that it involves just the notion of physical possibility; the essentialist could protest that this is an unnecessarily restrictive notion of possibility. After all an event's occurrence at t may be physically impossible relative to an earlier t yet nonetheless metaphysically

possible. But this point aside, what McGinn overlooks is that given that x does not exist **before** t_1, then there is no time t before t_1 at which x exists, so *a fortiori* there is no time t before t_1 at which x **exists** such that it was physically possible at t that x come into existence at t_1. So by principle O x's not-coming-into-existence at t_1 would be essential to x at t_1. But by hypothesis x comes into existence at t_1. Principle O must be rejected because it is contradictory.

As Michael Bradley has pointed out in a private communication, contradiction also arises in another way. Take any property F. Since x does not exist before t_1, there is no time t before t_1 at which x exists, so *a fortiori* there is no time t before t_1 at which x exists such that it was physically possible at t that x not be F. So by principle O x's being F would be essential to x. But F is any property. Let it be the property of being round and square. So by principle O x's being round and square is essential to it. But this is absurd. Principle O must be rejected.

At this point the origin-essentialist might change tack: before x exists we cannot refer to it; we cannot **refer** to future individuals. So that:

> before his coming to exist, Nixon was not a subject for the possibility that he should come to exist differently; there is no earlier situation containing Nixon with respect to which actuality might have turned out otherwise than it did (McGinn, p130).

Moreover, the notion of possibility operative here is that of **metaphysical** possibility. On this point compare Prior, who maintains that there can be no truths that are about any particular individual until there are such persons to be the subjects of these truths. Hence once an individual comes into existence, it is already too late for the individual in question to have any chance of having had other origins. (Prior 1959-60, p.695.) So on this view we cannot talk sensibly **before** x comes into existence of **x's** possibly having a different origin from what it will have. Doesn't this then secure the necessity of x's origin? No, because even if we

concede the impossibility of reference to x **before** it exists, surely we can refer to x once it exists or has existed, in which case why can't we say of x that it **could** have had a **different** origin?

The response might be made that an individual could have been different at t_1 only if it could have been, done, or suffered something before t_1 (Bogen, p.163.). Call this 'principle R'. Principle R appears to be a version of principle O. Bogen argues for principle R as follows:

> to make sense out of 'Napoleon need not have lost [the battle of Waterloo]' we must be able to explain how things might have been different before the time at which the defeat actually occurred (p.162). 'I don't mean that anything might have been different or might have happened differently before the defeat; all I mean is that Napoleon need not have lost the battle' is scarcely... intelligible (p.162). The reason that we can make sense out of 'Napoleon need not have lost' is that Napoleon existed before the defeat and was then capable of doing something or being affected in such a way as to prevent **his** defeat (p.163).

But since Nixon does not exist before t_1, **he** could not have been, done or suffered something before t_1 so as to affect his origination at t_1. Hence Nixon's coming into existence at t_1 could not have been different.

Bogen has in mind the notion of **physical** possibility. As such his argument has no bearing on the **metaphysical** possibility of Nixon's having had an alternative origin. Recall in section 3.2 that the constraints set by the state of an object at an earlier time on the states it **metaphysically** could have been in at a later time are far from restrictive. This contrasts of course with a situation involving **causal** possibility. But even given that causal possibility is the operative notion in Bogen's case, it is not clear that an opponent of origin essentialism must accept principle R. They can concede to Bogen that what would have happened after x comes into existence, would have happened too late to make x different at the time of its origination.. And certainly before he comes into existence Nixon cannot be, do or suffer anything, let alone something that will affect his future origination. But why must

Nixon be required to be, do or suffer something then in order to make sense of the **causal** possibility of his having had a different origin? Clearly as a result of some change in the world Nixon comes into existence. He suffers no change in coming into existence, but once he exists, changes that involve him are changes that **he** suffers. So if the change **in** the world which brought Nixon into existence is not a change in Nixon, why couldn't - **causally** couldn't - the change in the world which brought Nixon into existence have been **different**? *A fortiori*, why couldn't - **metaphysically** couldn't - the change in the world which brought Nixon into existence have been different?

3.6 MCGINN ON THE ORIGIN OF ORGANISMS

3.6.1 IDENTITY OF ZYGOTE WITH ADULT

While the arguments we have considered so far have primarily involved organisms, they apply equally *mutatis mutandis* to inanimate objects. We now turn to consider an argument by McGinn that seeks to establish the essentiality of origin for organisms as such. McGinn's argument is as follows. The adult (youth/child) organism is identical with its zygote. Since an object necessarily is identical to itself, then the adult organism, being an object, necessarily is identical with its zygote. The organism must have developed from the zygote it in fact developed from. Moreover, McGinn argues that the zygote necessarily developed from the gametes that actually fused to create it, so that an organism necessarily developed from the gametes it in fact developed from. Since, furthermore, these gametes in turn necessarily come from the parents they in fact came from, an organism could not have originated from parents other than its actual biological ones (this conclusion is of course not contradicted by the possibility of extra-uterine fertilisation or surrogate gestation).

McGinn's contention that an organism necessarily is identical with the zygote it actually develops from, rests on the claims that the organism is identical with its zygote, and that identity is necessary. McGinn assumes the necessity of identity rather than arguing for it. And as an essentialist he should accept it, for we recall from 2.2.1 that given the terms of the debate the essentialist should accept rather than reject this assumption. McGinn does argue for his second claim, viz. that the adult organism is identical with its zygote. Let us consider this argument.

Since McGinn conceives of organisms as continuants, it is clearly **possible** for the adult to be strictly identical with the zygote despite gross qualitative differences.

> [T]he claim that we have to do with a single persisting entity here [is not] out of the ordinary when one reflects on the drastic metamorphoses endured by seeds as they grow into trees and caterpillars as they become butterflies. (McGinn, p.132)

But is the adult **in fact** identical with the zygote? McGinn argues that it is, is so far as

> adults are commonly identical with children, and children with infants, infants with fetuses, and fetuses with zygotes. Any attempt to break the obvious biological continuity here would surely be arbitrary (McGinn, p.132).

Now typically we take spatio-temporal continuity to be a necessary condition of the identity of a continuant over time. But while necessary, spatio-temporal continuity is not sufficient. For clearly, we could trace the career of a car through a spatio-temporally continuous series of changes such that the car is instantaneously annihilated and replaced by a qualitatively identical car. Yet we would not standardly want to identify the car substituted with the car replaced. (According to some Catholic theologians we should make the identification, for the unity over time of a continuant consists in its perishing and being ever recreated by God at each new instant.). We also require that the successive stages in the history of a continuant, however spatio-temporally continuous, be **causally** related. (See Shoemaker 1979.)

But while necessary, spatio-temporal-**causal** continuity is not sufficient for identity over time, for while spatio-temporal- **causally** continuous, the series seed-grass-compost-peat-coal-diamond does not constitute the history of a **single** continuant. What we standardly require in addition is that such continuity be under an ultimate **sortal.** (See Hirsch 1982, chapter 2.) The theory of sortals and its essentialist implications will concern us in chapter 4. For the time being, we can roughly understand ultimate sortals to be those general terms like 'table', 'car', 'cat' and'human being', general terms which on one account - particularly the Wigginsian - are associated with principles for the individuation-at-a-time and the reidentification-over-time of an individual, which principles vary with the sortal in question. An essentialist theory of sortals requires that an individual falling under an ultimate sortal satisfy that sortal throughout its existence; it belongs essentially to that sortally determined kind[12] I shall examine this theory in chapter 4.

Ordinary language readily provides us with the sortals required to trace the career of an organism from zygotehood to adulthood. In the case of *homo sapiens* for example, the required sortal is 'human being'. The zygote is a human being. The adult is a human being. In fact the zygote is the very same human being as is the adult. They are identical. Being a zygote is a phase in the career of a single human being, as is being an adult. The terms 'zygote' and 'adult' are thus what Wiggins

---

12 The question why `red and white patch' is not thought to qualify as a sortal is discussed in chapter 4, as is the question why on the one hand grass/compost or coal/diamond is thought to mark a non-arbitrary discontinuity in the continuous series seed-grass-compost-peat-coal-diamond, while on the other hand seed-seedling-grass (but not seed seedling-grass-straw) is thought to constitute a non-arbitrary series (as is zygote-foetus-infant-child-adult, for that matter).

would call 'phase-sortals' (McGinn, p.132) rather than ultimate sortals. (The consideration that a zygote probably is not a person prompts McGinn to conclude that 'person' is not an ultimate but rather a phase-sortal - being a person is a distinct albeit long **phase** in the career of a human being (McGinn, p.132). We altogether ignore the vexed question of **personal** identity in this thesis). So if 'human being' is a sortal according to the Wigginsian account, then the adult will be identical with the zygote - and presuming identity to be necessary - necessarily identical with the zygote.

Forbes replies that an acceptation of 'zygote' as a phase-sortal would be idiosyncratic, in so far as in its standard **biological** use it denotes a certain natural kind of cell, and so is actually an ultimate sortal (Forbes 1985, p136), thus rendering it distinct from the adult, since on the theory of sortals objects falling under distinct ultimate sortals are not identical. Still, Forbes concedes that biological usage presents no final bar to a construal of 'zygote' as a phase-sortal. Admittedly, zygotehood is **unlike** the phase childhood in as much as unlike childhood it admits of very precise temporal boundaries (fusion-mitosis), but there would appear to be no *a priori* reason why phases must have vague temporal boundaries (Forbes 1985, p.126).

Forbes' real objection to a construal of 'zygote' as a phase sortal turns on the point that it forces a distinction between the individual organism and whatever constitutes it during its zygotehood. On this construal of 'zygote' a human for example cannot be identical with the sum of cells which constitute it at any moment (Forbes, p136). To see this, consider another phase, childhood: the child can be properly said to survive the replacement of cells that occurs in the transition to puberty. Even though the individual's **phase** as a child ceases, the child itself does not cease to exist. Since the child **survives** the supersession of cells, i.e. the child is identical

with the adolescent, the indiscernibility of identicals requires that the child be distinct from the cells that constitute it at any particular moment. And similarly for phases other than childhood; in particular, the zygote itself does not cease to exist, even when the cell divides. The **cell** itself **does** cease to exist when it divides - at least that is how we typically describe **amoebic** division, and with good reason, for upon division either the cell is identical with (a) both daughter cells or with (b) only one or with (c) neither. Given the symmetry and transitivity of identity, option (a) leads to the daughter cells' being identical, which is counter-intuitive. Option (b) raises the problem of which daughter cell to identify **non-arbitrarily** with the mother. So option (c) remains by elimination: the mother cell is identical to neither daughter cell because it ceases to exist upon division. So with mitosis the zygote does not cease to exist even though the cell that constitutes it does. But then the two are distinct. The upshot is that that adult itself will be identical with the zygote (just as the adult is identical with the child), but it will not be identical with the cell that constitutes the zygote. But as Forbes points out, it is precisely the cell - the particular fusion of sperm and ovum - which should count in McGinn's argument for the essentiality of an argument's origin: Given the necessity of identity:

> it will be ... necessary ... that the human is identical with the zygote he was - the human could not have been a different zygote, nor a different child - but for all that ... the human could have originated from a different **cell**, since a different cell could have made him up when he was a zygote, just as different cells could have made him up when he was a child. ...So ... if we use 'zygote' as an ultimate sortal for cells, then it is false that a human is identical to his zygote; but if we use it as a phase sortal for humans ... then although it is true that a human is (necessarily) identical with the zygote he was, this is insufficient to establish the [necessity of the human's coming from the cell which formed upon fusion of sperm and ovum] (Forbes 1985, p.137)

3.6.1.1 ZYGOTE AS DISTINCT FROM ITS CONSTITUTIVE CELL

But if 'zygote' really is a phase sortal for humans, the origin essentialist might reply, what is the difference between the **cell** formed by gamete fusion and the **zygote** this cell is supposed to constitute? Specifically, how could a **different** cell have made up the very same zygote? In reply one could point out that cells are continuants just as much as the organisms they are part of, and that through metabolic processes the material composition of a cell is changing continually. Since this is no bar to our identifying a cell after some bio-chemical reaction as the same cell as a cell before the reaction, why shouldn't we be allowed to identify a certain counterfactual cell c* which has the same career as an actual cell c but differs at the beginning of its career only slightly in material composition from c, as that very cell c? And since c constitutes zygote z, then c* could have constituted z. What does press for an answer is the question how much c* could have diverged in both material composition and genetic structure from c.

Taking the issue of genetic structure first. Since biological development is a function of genetic structure among other things, c* would presumably have had to have a genetic structure closely similar to c's actual structure, to provide for an evolution from zygote to adult that would have paralleled z's actual evolution to adult a. So is having a genetic stucture that is closely similar to the actual genetic structure of the zygote cell essential to the organism o that develops from the zygote? In other words, could o have developed from a zygote with a **more than slightly** different genetic structure? It would appear that the essentialist should allow the possibility, given that they standardly allow an organism a wide range of possible divergence from its actual development. But then there seems no reason why o could not have developed from a zygote with a **substantially** different genetic structure. After all, consider the possibility that my actual zygote had been

subject to radiation that induced genetic mutation. Nonetheless could I not have been the organism that developed therefrom, regardless of the extent of mutation?

Turning now to the issue of material composition. Suppose that my actual zygote had been composed of completely different matter. Perhaps it was piecemeal and completely replaced with bits of matter, just like the ship of Theseus, and that the original matter had been dispersed throughout the universe, so that there is no complication of competing candidates (barring scattered objects). Suppose furthermore that the ensuing zygote was genetically identical to my actual one. Could I have been that zygote? Suppose in addition that this zygote underwent radioactively induced genetic mutation. Could I have been the adult that eventually evolved? If affirmative replies to these questions are forthcoming, then not only need an organism not have developed from the zygote cell it in fact developed from, but it need not have developed from a cell with even **substantially** the same genetic structure or material composition as its actual zygote cell.

The origin essentialist could choose to reply that a zygote cell could not have been composed of completely different matter, i.e. of no molecules in common with its actual molecules. Similarly they might insist that the zygote cell could not have had a genetic structure completely dissimilar from its actual structure. But then, as we shall presently see, the origin essentialist affirms the possibility of the zygote cell's having had a **slightly** different material composition (or slightly dissimilar genetic structure) on pain of contradiction, at least on the assumption that the accessibility relation between possible worlds is transitive. And the denial of **this** assumption makes for an unintelligible metaphysics, even if we grant the coherency of the essential/accidental distinction. So consistency requires that the impossibility of **complete** difference go along with the impossibility of **any** difference, however slight. By embracing consistency an origin essentialist will then have rescued their

position, but at the cost of alienating other origin essentialists who endorse the possibility of **some** difference. I shall now substantiate these claims.

3.6.1.2 THE POSSIBILITY OF SLIGHT DIFFERENCE IS INCOMPATIBLE WITH THE IMPOSSIBILITY OF COMPLETE DIFFERENCE

3.6.1.2.1 THE CASE OF MATERIAL COMPOSITION

To the task to hand I shall freely adapt a discussion by Carter, recasting his references to possible worlds in terms of ordinary language 'could have' and 'would have'. (see Carter 1983, pp.225-229.) Suppose that a cell c - which is notionally but not physically subdivided into one hundred arbitrary portions of equal mass p_1, p_2, ..., p_{100} - could have come into existence with, say, 98 % of its actual material composition m, the remaining 2 % being distinct from any portion of m. So c could have been constituted by m*: p_3, p_4, ..., p_{100}, p_{101}, p_{102}. Equally c could have been constituted by m#: p_1, p_2, p_5, p_6, p_{100}, p_{103}, p_{104}. So the cell which could have been composed of m* could have also been composed of m#. Hence the cell which **would** have been composed of m*, had things turned out otherwise, could have been composed of m#, and conversely, the cell which **would** have been composed of m#, had things turned out otherwise, could have been composed of m*. But in both of these counterfactual cases the matter which would have been common to m* and m# would have been less than 98 % of m. So the cell which would have been composed of m* (or m#), had things turned out otherwise, could have been composed of less than 98 % of m. But the cell which would have been composed of m* (or m#), had things turned out otherwise, is none other than c. So if c could have been composed of at most 98 % of m, then it could also have been composed of **less** than 98 % of m. By parity of reasoning, given a threshold of 1/n th, where 'n' denotes the total number of **fundamental** and thus indivisible particles

comprising m, then c could have been composed of **none** of m's n portions, if, that is, c could have been composed of at most 1/n th of m (i.e.of at most one of m's portions).

Carter (Carter 1983, pp.227-228) explains that one who wanted to reject the conclusion of this argument might deny that the cell that would have been composed of m* (or m#), had things turned out differently, **could** have been composed of m# (or m*). That is, in the technical idiom of modal logic, a world in which c would have been composed of m# is **inaccessible** from (i.e.not possible relative to) a world in which c is composed from m*, and conversely. Letting 'Axy' mean 'x is accessible from y", 'm_c' mean 'the possibility of c's being composed of m', 'm_c*' mean 'the possibility of c's being composed of m*' and 'm_c#' mean 'the possibility of c's being composed of m#', we have it that Am_c*m_c and $Am_c\#m_c$ but not that either $Am_c\#m_c*$ or $Am_c*m_c\#$. But if accessibility is an equivalence relation, then the conjunction of 'Am_c*m_c' and '$Am_c\#m_c$' entails that $Am_c\#$ m_c* and $Am_c*m_c\#$. So accessibility cannot be an equivalence relation; in particular it would be held to be intransitive, presumably because the denial of symmetry is thought to be more radical than the denial of transitivity.

In section 3.7.3.2 I shall argue to be bad metaphysics the notion that something's being possible depends on what is taken to be actual. But even if we grant the coherency of the notion, it is not clear that neither '$Am_c\#m_c*$' nor '$Am_c*m_c\#$' holds. In particular, the essentialist would surely grant the possibility of a certain cell's being composed of m#. Suppose furthermore that the cell in question, call it 'C', belonged to a world W whose material composition and configuration duplicated precisely at all times the world w# in which $m_c\#$ obtained. Carter urges that in that case W **is** w#. In so doing he appeals to what I have called 'PIOICP', the Principle of the Identity of the Qualitatively Indiscernible in Cases of Counterfactual

Possibility. Furthermore he argues that since W is surely possible from the vantage point of the world w* in which m_c* obtains, then w# is possible relative to w*. (see Carter 1983, pp.228-229). In other words $Am_c\#m_c$*. Similarly, this argument can be rerun to show $Am_c*m_c\#$.

It has been said that there is nothing as dangerous as a philosopher in the grip of a theory. Perhaps the danger lies mainly for the truth. In any case Salmon appears to confirm this adage. For while he at one stage appears to grant PIOICP in the form pertinent here, namely that names which denote materially indiscernible worlds are coreferential, in response to Carter he explains that he rejects this interpretation of his view. (Salmon 1982, pp.237-238.) What he endorses is rather the principle that **mutually accessible** worlds, which are materially indiscernible, are identical. And presumably what determines whether worlds are mutually accessible is not mere material indiscernibility but **factual** indiscernibility, i.e.whether the same facts obtain in those worlds, including facts about the identity of the inhabitants of those worlds. (see Salmon 1984, p.116.) I find this position most unsatisfactory. On what are facts about the identity of objects between worlds to supervene, given that they are not to supervene on mere qualitative indiscernibility? Salmon's response appears a desperate attempt to save un untenable position, namely that accessibility is not transitive. (This point will be developed further in 3.7.3.2.) Since this position does appear untenable, we ought to conclude that the impossibility of complete dissimilarity is inconsistent with the possibility of some dissimilarity of the cell's material composition.

3.6.1.2.2 THE CASE OF GENETIC STRUCTURE

Similarly, the argument can be rerun to show the inconsistency of the impossibility of complete dissimilarity of the cell's genetic structure with the possibility of some dissimilarity. This of course holds providing that degrees of similarity or dissimilarity of genetic structure can be determined. This is not a question - as it is in the case of material composition - of the extent of overlap between worlds of numerically identical molecules. Suppose we can overcome this problem. It would follow that the denial of the possibility of complete dissimilarity of genetic structure requires that the cell have the very genetic structure it actually has. It does not follow from this however that the cell could not have had a widely divergent career, for while biological development is a function of genetic structure, it is not wholly so; indeed consider the case of identical twins. If on the other hand, it is allowed that the cell could have had a slightly different genetic structure, then it has to be allowed that the cell could have had a completely different genetic structure. In that case the essentialist will have to grant *contra* the sortal esentialist that a human cell could have been a frog's cell, and thus that a human being could have been a frog. To thwart this conclusion the sortal essentialist will have to insist that a cell could not have had even a slightly different genetic structure. I address the plausibility of this position in chapter 4.

In any case the impossibility of complete difference (i.e. dissimilarity) of **material composition** requires the impossibility of any difference of material composition. There then arises the question of the plausibility of the thesis of the impossibility of complete difference of material composition. I consider this issue in a more general setting in section 3.7.

3.6.2 GAMETES AND THEIR PARENTS

Having examined McGinn's argument for the claim that an organism came essentially from the zygote cell it in fact came from, I now turn to examine (a) McGinn's contention that the zygote cell itself was essentially formed from the gametes it in fact fused from, and (b) his contention that these gametes in turn essentially came from the parents they in fact came from.

Let us first examine his claim that none but the parent that actually produced the gamete **could** have produced it. Here McGinn's reasoning is vague: he seems to appeal to the consideration of the continuity of the gamete with the parent. But here continuity does not seem to secure the essentialist conclusion, for, as Elliot and Gallois observe:

> the particles which constituted the gamete provided by the Queen's actual father could have become incorporated [through ingestion, inhalation, or injection] in Mr. Truman [instead of in King George], and could have arranged themselves in exactly the way in which they were arranged in the Queen's actual father. A parallel story could have been true of the particles forming the gametes provided by the Queen's actual mother, finding their way into Mrs. Truman's body ... . Mr and Mrs Truman copulate and a zygote forms from the gametes just described. The zygote develops in the normal way and a child is born. (Elliot and Gallois, p.292).

Perhaps the story is exceedingly far fetched. Nevertheless the essentialist would presumably grant that it at least seems to be metaphysically possible. If so, wouldn't the zygote (and the child) have been the Queen, given that the gametes were the ones she actually developed from? But her parents would then have been the Trumans. If so, it would not be essential to the Queen's gametes, and thus not essential to the Queen either, that she had the parents she **in fact** had.

3.6.3 THE ESSENTIALITY OF GAMETES

Let us turn now to consider McGinn's claim that one had to come from the gametes one in fact came from. The reasoning is as follows:

Just as you must have come from the zygote you came from because you are ... developmentally continuous with it, so you must have come from the gametes you came from because you are similarly continuous with them (McGinn, p.133).

McGinn thinks the gametes cannot be identical with the zygote (McGinn, p132) presumably because two things a and b cannot both be distinct from each other yet identical to a third thing, given the transitivity and symmetry of identity. (An alternative explanation of the non-identity of gametes with zygote woul;d characterise the transition from gametes to zygote as a non-arbitrary (i.e. natural) discontinuity, marked by the application of different ultimate sortals on either side of the breach, so that while human gametes as opposed to feline gametes may be **human**, they do not fall under the **sortal** 'human being', unlike a human zygote. But this alternative explanation would have to explain why the gamete-zygote transition is sortally discontinuous even when spatio-temporal-causally continuous. Again a theory of sortals would have to explain this). In any case does mere continuity from gametes to zygote suffice for the essentiality of gametes? For could not alternative gametes have fused to form the actual zygote? If we treat 'zygote' as an ultimate sortal, then the zygote is identical with the cell (for then it is a **kind** of cell); if we treat 'zygote' as a phase sortal then the zygote is merely constituted by the cell. In either case, could any of these cells have formed by the fusion of alternative gametes? And how different - if at all - could such gametes have been from the actual gametes? After all gametes are cells, and cells can persist through metabolic changes (and thus changes of material composition). They are thus continuants that can **survive** a loss of original material parts. So could they have

come into existence with slightly different or altogether different material parts? If the answer is 'no', the zygote develops essentially from its actual gametes. If the answer is 'yes', then obviously it does not develop essentially from its actual gametes. This question leads us directly to the issue of the alleged essentiality of original material composition.

3.7 KRIPKEAN ARGUMENTS FOR THE ESSENTIALITY OF ORIGINAL MATERIAL COMPOSITION

3.7.1 KRIPKE'S ARGUMENT

Kripke's own argument is stated as follows:

> *If a material object has its origin from a certain hunk of matter, it could not have had its origin in any other[hunk of] matter [either of the same kind or of a different kind]* ... Let 'B' be a name (rigid designator) of a table, let 'A' name the piece of wood from which it actually came. Let 'C' name another piece of wood. Then suppose B ...[is] made from A, as in the actual world, but also [that] another table D ...[is] simultaneously made from C. (We assume that there is no relation between A and C which makes the possibility of making a table from one dependent on the possibility of making a table from the other.) Now in this situation $B \neq D$; hence, even if D were made by itself, and no table were made from A, D would not be [table] B. [Since nothing in this argument turns on B's or D's being a table, it is perfectly general and thus applies to any material object, artifactual or non-artifactual, animate or inanimate.] (Kripke, p.114.)

What this argument is supposed to establish is that table B **could** not have **originated** from hunk C, but given the necessity of distinctness, all it in fact establishes is that **in the circumstances** of table D's being made on its own, D **would** not have **been** table B. (See Salmon 1982, p. 204.) So in **these** circumstances B **would** not have been made from C. But this as it stands is insufficient to show that B would not have been made from C in any **other** circumstances, in particular in circumstances where D is made from something other than C. And clearly Kripke has given us no reason to believe that D couldn't have been made from something other than C. Hence Kripke's argument fails to show

that B **could** not have been made from C. So Kripke has not given us any reason to believe that B could not have had an origin other than its actual origin A.

Salmon steps into the breach with two arguments which he thinks secure Kripke's desired essentialist conclusion. (In what follows read Salmon's 'it is possible that' as 'it was or could have been the case that'.)

3.7.2 SALMON'S ARGUMENT 1

My discussion will be brief, as this argument is jejune, indeed verging on the egregious. The argument proceeds as follows:

> [Assume principle (III):] For any possible tables x and x*, and any possible hunks of matter y and y*, if it is possible for table x to be originally constructed from hunk y while hunk y* does not overlap with hunk y, and it is also possible for table x* to be originally constructed from hunk y*, then it is also possible that table x be originally constructed from hunk y ...[while simultaneously] table x* be originally constructed from hunk y*. ... Principle (III) trivially entails [by contraposition], ... letting x and x* be the same, that if it is merely possible for a given table x to originate from a certain hunk of matter y, then it is in fact necessary that the given table does not originate from any nonoverlapping hunk of matter y*, *provided that* it is impossible for the given table x to be constructed entirely from all of hunk y and ...[simultaneously] entirely from all of the distinct hunk y*. (Salmon 1982, p.213)

While valid, Argument 1 trips up on its premise principle (III), whose adoption renders Argument 1 either unsound or circular. The impossibility referred to in the proviso of the essentialist conclusion obtains only if the variables 'y' and 'y*' have distinct entities as their respective values. But then the variables 'x' and 'x*' must have **distinct** values when the variables 'y' and 'y*' have **distinct** values, lest Argument 1 become either unsound or circular. For in the cases where 'x' and 'x*' refer to the **same** entity when 'y' and 'y*' refer to distinct entities, the consequent of premise (III) will be false. Now while the consequent is false, the antecedent will come out either true or false. But if it comes out **true** then (III) as a whole will be false and so Argument 1 will be **unsound**. And if the antecedent comes out **false**

then (III) will come out true - at least in classical logic - but then at least one of the conjuncts of the antecedent of premise (III) will be false. That is to say that it will be false that it is possible for a given table to be constructed from a certain hunk. In other words it will be impossible for that table to be constructed from the hunk. But what licenses **this** claim? Isn't that just what Argument 1 is trying to establish? Thus Argument 1 will be **circular** if the antecedent of (III) is false when its consequent is false. Hence if 'x' and 'x*' refer to the same entity then premise (III) will be either false or assume the very point at issue, thus rendering Argument 1 either unsound or circular. So if the argument is to be neither unsound nor circular then (III) will have to be understood as requiring that the variables 'x' and 'x*' have distinct values. Since Argument 2 contravenes this requirement it is either unsound or circular.

3.7.3 SALMON'S ARGUMENT 2

I said that Salmon had two arguments for the essentiality of material origin. We now turn to consider the second argument.

Definition 1: To say that a table x was **originally constructed** (originated) from a hunk of matter y, is to say that **every** original part of x came from y, and that **every** portion of y contributed to forming x.

Premise IV*: For any possible table x and any possible hunks of matter y and y*, and any plan P, if it is possible for x to be originally constructed from y according to plan P, where y* **does not overlap** with y, and it is also possible for a table to be originally constructed from y* according to plan P, then it is also possible [both] that x be originally constructed from y according to P, ... [while] in addition ...that [simultaneously] some table or other be originally constructed from y* according to P. (The provision that the tables be constructed **according to the same plan** is

added in response to the consideration that it is not clear that the table which would have resulted from a rearrangement (i.e. interchange and/or displacement) of the constitutive molecules of a table T would have been identical with T.)

Premise I: It is impossible for the same table x to be originally constructed from a hunk of matter y and in addition to be simultaneously originally constructed from a distinct hunk of matter y*.

Premise V*: If it is possible for a table x to be originally constructed from a hunk of matter y according to a certain plan P, then necessarily, any table originally constructed from hunk y according to the same plan is the very table x and no other.

(In what follows all references to the **construction** of a table or to its **origination** are to be understood as references to its construction or origination **according to a plan P**. More generally, in the case of non-artefacts such as mountains or trees one can understand 'to be constructed' in the looser sense of 'to be formed'. Thus the argument can be extended to cover non-artefacts.)

> Let w_1 be some possible world [not necessarily the actual] in which an arbitrary table B originates from some hunk [i.e. portion, sample, bit] of matter A. Let C be any hunk of matter that does not overlap with A in w_1 ... Suppose now that it is possible to construct a table from hunk C. Then by premise ... [(IV*)], there is...a possible world w_2 in which ... B originates from hunk A, just as in w_1, but also in which a second table, which we shall call 'D', originates from hunk C. By premise (I), tables B and D are distinct in w_1. It follows by the necessity of identity and distinctness that ... B and D are distinct in every possible world. Now consider an arbitrary possible world w_3 in which some table is constructed from hunk C. Could that table be the very table B from w_1? Given premise ... [(V*)], it cannot. For by premise ... [(V*)], the table in question in w_3 is none other than table D, and B and D are distinct entities in every possible world, including w_3. Therefore, there is no possible world in which Table B originates from hunk C. [Hence, if it is possible for a given table to originate from a certain hunk of matter according to a plan P, then it is necessary that the given table does not originate from any **nonoverlapping** hunk of matter. Thus B's material origin is not only essential to it but uniquely essential; nothing else could have had it, either accidentally or essentially.] (Salmon 1982, pp.206 -207).

It should be emphasised that this essentialist claim applies only to cases where the alternative origin does **not overlap** with the actual matter. Nothing follows about cases where there **is** some such overlap. This issue will be addressed later. Salmon calls premise (IV*) (or at least an analogue thereof) 'trivial and uncontroversial enough for the purposes at hand' (Salmon 1982, p.203). He similarly finds premise (I) 'philosophically uncontroversial' (Salmon 1982, p.200) and premise (V*) 'exceedingly plausible, almost to the point of being indubitable' (Salmon 1982, p.211), and says

> If two tables in two different possible worlds are constructed from the very same stuff in precisely the same way and ... with exactly the same structure atom for atom, how can they fail to be the very same table? What more could one ask? What more is there to being *this very table?* (Salmon 1982, p.211)

Nonetheless Salmon ends up revising premise (V*), replacing it with the weaker premise **(V**)**: if it is possible for a table x to be such that it is originally constructed from a certain hunk of matter y according to a certain plan P, **and** such that x is the **only** table so constructed, then necessarily, any table that is the only table to be originally constructed from y according to P is identical with x.

Premise V** together with premise (I) and a correspondingly strengthened version of premise (IV*) constitute what I shall call **'Salmon's modified argument'**. From Salmon's modified argument there follow two essentialist theses, which are weaker than Salmon's original conclusion that if it is possible for a given table to originate from a certain hunk of matter according to a plan P, then it is necessary that the given table does not originate from any nonoverlapping hunk of matter according to that plan. The first essentialist conclusion that follows is that if it was or could have been the case that a table x was such that it was the only table originally constructed from a hunk of matter y according to plan P, then x could not have been originally constructed from any **nonoverlapping** hunk y*. The second essentialist conclusion

follows by contraposition from the first conclusion: if it was or could have been the case that a table x was originally constructed from y, then x could not have been the only table originally constructed from any **nonoverlapping** y*. (Salmon 1982, p.229)

But why does Salmon replace premise (V*) with (V**)? Salmon's explanation is that (V*) conflicts 'with the plausible view that a given object can maintain its identity through time by an appropriate sort of spatiotemporal continuity though its matter is frequently replenished.' (Salmon 1982, pp.228-229) In particular, he cites Chandler's case of a ship s that is made from material m which is gradually replaced in the manner of the ship of Theseus until it no longer comprises s but lies in a disassembled heap. Call the ship that is still afloat in 1996 while m lies in a heap 's_c'. Suppose also that m is reassembled in 1996 to resemble s exactly as it was when first built. Call the reconstructed ship 's_r'. According to a dominant line of thinking s_c has a stronger claim to identity with s than does s_r. If so then we have **distinct** objects s and s_r with the **same** origin, which contradicts (V*). Suppose that in response (V*) is modified to be understood as governing cases of **transworld** identity only and not intraworld identity. Then (V*) so understood runs foul of the metaphysical if not causal possibility that s could have survived intact over the millenia without any change in material composition. Call such s in 1996 's*'. Is s* identical with s or with s_r? Or speaking more strictly, would s* have been s or s_r or neither? As Noonan observes: 'However one answers one has to abandon the claim that sameness of origin is a sufficient condition of trans-world identity, [i.e.that (V*) is true]' (Noonan, p.4).

Salmon's response to this problem is to replace (V*) with (V**).

(V)**: if it is possible for x to be the **sole** F to be originally constructed from a hunk of matter y according to plan P, then necessarily, any F that is the sole F to be

originally constucted from y according to P is identical with x.

3.7.3.1. A PROBLEM FOR V**

Gregory Currie has suggested in discussion that works of art represent clear counterexamples to (V**). He suggests that the very same abstract configuration of the very same paints on the very same canvas could have been a distinct painting from the painting it actually is, had its painter intended the configuration to represent something other than it was actually intended to represent. I am not sure what to make of this suggestion. If valid, it would follow that the same lump of naturally formed and aesthetically pleasing stone discovered in a quarry and subsequently exhibited in a gallery, could have been distinct from itself, if the discoverer of the rock had decided to have the rock represent something other than they actually decided. But this seems absurd. In short, I do not think that the identity of a work of art can be determined even partly by representational intentions. This presumably makes me an internalist as opposed to an externalist about the identity conditions of representational artifacts.

Works of art aside, Noonan faults (V**) on the following grounds. The essentialist ought to allow that it could have been the case that (1) while s_c existed in 1996 its original matter m was not reassembled but burnt. Furthermore the essentialist ought to allow that it could have been the case that (2) s was never built, so that s_c did not exist in 1996, yet m was assembled (not reassembled) into a ship in 1996 at the time s_r was. The ship would have resembled s_r exactly and would also have had the subsequent history of s_r. In each of these two cases there would have been only one ship in existence, so that according to (V**), the ship in the first case would have been the ship in the second case. Yet it would seem that the ship in the second case would have been correctly describable as s_r; after all, the ship would have had the same original matter, moment of origin and history as s_r. (This is of course to

appeal to something like PIOICP, the Principle of the Identity of the Qualitatively Indiscernible in Cases of Counterfactual Possibility.) And by appeal to the same sorts of considerations we are licensed to refer - as we did - to the ship in the first case as s_c. But s_c and s_r are distinct ships, hence the ship in the first case is distinct from the ship in the second case, contrary to (V**). (Noonan, p.6)

Noonan's own proposal, which he submits after dismissing several others (Noonan, pp.6-8), is that sameness of original matter plus sameness of thing-kind (which, although Noonan does not point this out, would have to include design) **plus sameness of moment of origin** is sufficient for transworld identity. (Noonan, p.4)

Noonan's Proposal: if it is possible for x to be the sole F to be originally constructed from a hunk of matter y according to plan P at a certain time t, then necessarily any F originally constructed from y at t according to P is x.

(Although Salmon does include references to moment of origin in more developed formulations of (V**), this is done without explicit motivation; hence I shall call such amended formulations of (V**) '**Noonan's** proposal'.) What I find interesting as an anti-full-essentialist is the question how one might defend Noonan's proposal. Noonan observes that one might try to ground it on a weaker condition of transworld identity, which we shall call '**WC**'. WC is the claim that sameness of **all but a sufficiently small portion** of original matter plus sameness of thing kind plus sameness of moment of origin is sufficient for transworld identity. If WC were true then Noonan's proposal would follow *a fortiori*. But according to Noonan WC is not true, for it is self-contradictory. (Noonan, p.8) (Noonan thinks his own proposal to be consistent. (Noonan, p.1)) In fact Noonan is confused; all that follows from his discussion is that WC is inconsistent with the assumptions of Salmon's modified argument for the essentiality of original matter. (Salmon's **modified** argument for the essentiality of original matter turns on the necessity of

distinctness **and** Noonan's proposal). So all that Noonan's discusion shows is that a commitment to the necessity of distinctness (LNI) entails that one cannot consistently hold both Noonan's proposal **and** its weaker version WC. Therefore one cannot argue **for** Noonan's proposal **from** WC, given a commitment to LNI. On this score Noonan is correct. (Moreover, as we shall see shortly, Salmon argues that WC is in fact false.)

But now to expand: suppose without loss of generality that identity of, say, 75 per cent of original matter plus identity of moment of origin and thing kind is sufficient for transworld identity. Suppose further that a table t is actually made from parts 1, 2, 3 and 4. Then since a table could have been made from parts 1,2,3 and 5 - where 5 is wholly distinct from (i.e. non-overlapping with) 4 - at the same time as t and according to the same plan, this table (call it 't*') would have been t, according to WC. So t could have originated from a distinct hunk of matter which overlaps with its actual original matter. (As it stands this does not contradict Salmon's essentialist conclusion, which asserts merely that an object could not have originated from a hunk of matter that does **not overlap** with its actual original matter.) Now a table could have been made under the same circumstances as t* from 1,2,5 and 7, where 7 is wholly distinct from 1,2,3 and 5. Call this table 't**'. So by WC t**would have been t*. Equally a table could have been made from parts 2, 3, 4 and 6, where 6 is wholly distinct from 1-4, at the same time as t and according to the same plan. Call this table 't#'. By WC t# would have been t. And a table could have been made under the same circumstances as t# from 3,4,6 and 8, where 8 is wholly distinct from 2,3,4 and 6. Call this table 't##'. By WC t## would have been t#. Now by Salmon's modified argument for the essentiality of original matter - which argument we recall comprises premise (V**), a correspondingly strengthened version of (IV*) and premise (I) - it follows that t** would not have been t##, for the original matter of one would not have overlapped with the other. So by the

transitivity of identity t* would not have been t##. Hence, again by the transitivity of identity, t* would not have been t#. In other words t* would have been distinct from t#. But by WC t would have been t* and t would have been t#. Therefore by the symmetry and transitivity of identity t* would have been t#, which contradicts the conclusion that t* would have been distinct from t#. (Noonan, pp.11-12). Thus given a commitment to identity's being an equivalence relation, to the necessity of distinctness and to Noonan's proposal, one must reject WC, i.e. the essentialist cannot consistently claim **both** (1) that any object originating from **most but not all** of a's original matter would have been a **and** (2) that any object originating from **all** of a's original matter would have been a.

3.7.3.2. A PROBLEM FOR NOONAN'S PROPOSAL

As for Noonan's proposal, there is a separate *ad hominem* bearing on **it**, which Salmon dubs 'the four worlds paradox'. Again for the sake of simplicity and without loss of generality let us take the case of our table t made from parts 1,2,3 and 4. Assume A: a material object could have originated from at least three quarters of its actual original material. Salmon finds A 'plausible'. (Salmon 1982, p.234. Salmon's threshold is actually 98% rather than 75%) Note that to assume A is not to assume WC. Indeed Salmon argues that WC is in fact false, i.e. he argues that there are possible worlds containing a table with at least three quarters of, say, t's original matter, which table is nonetheless distinct from t. But to secure this conclusion Salmon needs a further assumption A*, which he assumes without acknowledgement: a material object cannot originate from less than three quarters of its actual original material. Without A* he cannot justify his claim that a table built from less than this proportion is not t. (Salmon 1982, p.231) So since it is possible that a table originate from 1,2,5 and 6, this table cannot be t. Call this distinct table 'T'. By assumption A table T could have originated from 1,2,4 and 5.

But by the same assumption **t** could have originated from 1,2,4 and 5. (Salmon 1982, pp.230-231) So distinct tables t and T could have originated from the same matter, contradicting Noonan's proposal and *a fortiori* WC. Hence, since nothing hinged on A's and A*'s threshold's being seventy five per cent rather than any other arbitrary figure, then given any pair of principles like A and A*, Noonan's proposal must be abandoned, at least if the accessibility relation between possible worlds is transitive.

Salmon wants to keep Noonan's proposal since he wants to argue for the essentiality of material origin. But rather than abandon either of A or A* Salmon chooses to deny a hitherto implicit assumption, viz. the transitivity of the accessibility relation between possible worlds, which is thus to abandon systems like S_4 and S_5. He would then have to amend Noonan's proposal thus:

Amended Version of Noonan's Proposal:

if in a world W accessible from a world w, x is the only F to be originally constructed from a hunk of matter y according to plan P at time t, then in any world accessible from w any F originally constructed from y at t according to P is x.

Salmon can then argue, as he does, that while table T which is actually built from 1,2,5 and 6, could have been built from 1,2,4 and 5, this possibility is of second degree only; i.e as things actually stand it is **not** possible that T should have been built from 1,2,4 and 5, but it **would have** been possible had things turned out differently. So while there are worlds in which T is built from 1,2,4 and 5, these worlds are inaccessible from the worlds in which t was built from 1-4. (Salmon 1982, pp.238-240) So since the conditions for the application of an amended version of Noonan's proposal are not met in the case of tables t and T, we do not have a genuine counter-example to it.

In asserting that T could have been built from 1,2,4 and 5, Salmon's strategy requires that the worlds in which T is so built are inaccessible from all the worlds in which t is built from 1,2,3 and 4. This means that it is **actually** (i.e. from the point of view of the actual world) impossible that T should have been built from 1,2,4 and 5, but that had things turned out otherwise, it would have been possible. As Carter points out, the effect of this approach is to relativise essential properties to worlds: an object's essence, the sum of its essential properties, is contingent on what actually occurs. (Carter 1983, p.227.) Thus for example had t been originally constructed from 1,2,4 and 6, it would not have been possible for t to have been constructed from, say, 2,3,4 and 5 (given A and A* of course). So while it is **not** in fact essential to t that it was not originally constructed from 2,3, 4 and 5, this **would have** been essential to t, had it been built from 1,2,4 and 6. (Cf. Carter 1983, pp.227-228.)

I have two reservations about Salmon's strategy.

Firstly, it is not clear that the worlds in which T is built from 1,2,4 and 5 are - as Salmon requires - inaccessible from the worlds in which t is built from 1,2,3 and 4. After all, given that t is actually made from 1,2,3 and 4, why could it not have been the case that a table **distinct** from t was built from 1,2,4 and 5? Salmon takes it that the truth of an amended version of Noonan's proposal simply guarantees the impossibility of this. (Salmon 1982, p.240.) Clearly this is unsatisfactory.

Secondly, just how can something be actually impossible but possibly possible? That is to say, how can it be the case that while T could not have been built from 1,2,4 and 5, given that t was actually built from 1,2,3 and 4, still had t been built from 1,2,3 and 6, T could have been built from 1,2,4 and 5? Moreover, why is it **absolutely** impossible, i.e impossible relative to **any** world, that a certain table t originate from less than three quarters of its actual original material, when in

contrast the possibility that t originate from certain material m is not absolute but relative only to **certain** worlds? This absolute impossibility is required to secure relative or conditional possibility. But then this absolute possibility appears arbitrary or at least unmotivated.

Now as Carter observes, conditional possibilities are intelligible with reference to causal possibility, but not to metaphysical possibility. (Carter 1983, p.223.) After all, how could what **actually** happened determine what could have happened and what could not have happened, when **before** it actually happened it too, like what did not happen, was nonactual? After it had happened it would surely have been too late to determine what could have happened. So similarly, how could what **merely could have** happened but did not happen, determine what could have happened or not, when before its possible occurrence it too was nonactual? If something could **not** have happened it would not be **just** because something else did happen, for even things that **could** have happened do not happen because something else did happen. Rather the thing that did happen, happened because the things that did not happen, as it turns out, did not happen. In short I find suspect the notion of conditional possibilities which underlies Salmon's rejection of the transitivity of the accessibility relation. Many essentialists would probably agree with me. Again I suppose many would not. Thus the latter might well support Salmon's attempt to reconcile assumptions A and A* with an amended version of Noonan's proposal.

3.7.3.3 EVALUATING NOONAN'S PROPOSAL

The upshot of my discussion of Salmon's second argument for the essentiality of origin is that its key premise, Noonan's proposal, with or without A or A*, remains unsubstantiated. Is there any consideration then which weighs in its favour? Otherwise Salmon's modified argument for the essentiality of original matter will be unfounded in as much as its assumption of Noonan's proposal is unfounded. After

all, what is so special about **original material composition**, when it is granted that an object could have had a **career** that is radically different from the one it actually had? Now Noonan's proposal is equivalent to the claim that no object other than a can have a's original matter, i.e. it is essential to every object distinct from a that it not originate from a's original matter. Hence Noonan's proposal is itself an essentialist thesis of sorts. (cf. Noonan, p.3.) All that seems to ground it is intuition; indeed Noonan calls it 'the commonsense point of view'. (Noonan, p.8.) But this intuition is not universally shared, even among non-Leibnizians. We have just seen how a committed full-essentialist like Salmon ends up **qualifying** this intuition in the light of his rejection of the transitivity of modal accessibility between possible worlds. And of course there are anti-full-essentialists like Penelope Mackie who would reject altogether the intuition under discussion, even though she accepts the idea that things could have been otherwise. Since I carry no brief for essentialism, let alone full-essentialism, it is not for me to adjudicate between these conflicting intuitions. The upshot is that a Kripkean argument like Salmon's for the essentiality of original matter remains deadlocked over indecision about the transworld sufficiency condition I dubbed 'Noonan's proposal'.

3.8 AN ARGUMENT AGAINST THE ESSENTIALITY OF ORIGIN

For her part Marjorie Price is one who accepts the notion that things could have been otherwise but provides an argument **against** the essentiality of origin, urging that in fact the commonsense point of view is on **her** side. I cite her argument at length:

> ... one can easily imagine a material object's having its origin in stuff different from that which originally constituted it. Suppose that in the actual world w Jones makes table B during time-stretch t_1 - t_2 out of the wood in block of wood A. Now consider a possible world w_4 whose history diverges from w's at t_1. In w_4, Jones at t_1 splits A into a piece of wood C, whose matter is seven eighths by weight of A's matter, and a piece E consisting of the remainder of the wood that constitutes A. He then uses C's wood to build a table D, which resembles B in all external details, except, of course, for being proportionately

smaller than B.

It seems to me that we would not say that, despite its similarities to B, D is a different thing. We would view w_4 as a world in which Jones constructs B with less wood than he uses to make B in w. Surely it is possible for B, that very table, to have been more parsimoniously made than it actually was.

An advocate of the thesis that origin is necessary might try to counter this criticism in the following way: he will describe another possible world w_5 in which Jones uses the wood in block A to fabricate a table F, which resembles B in all external details. Jones gradually replaces F's parts and later builds D out of seven-eighths by weight of the wood that constitutes F's original parts.

We would not, the rejoinder will proceed, decidedly prefer D to F in identifying a table in w_5 with B. This shows that you really were not entitled to assert that D in w_4 is the same thing as B. Moreover, since ... [a] physical thing could have been made from a different articulated object than the one it had its origin in, [- thus e.g. a medallion could have been made from a crucifix instead of a ring as long as the constitutive gold had been the same -] you cannot protest the stipulation that the smaller table in w_5 is ... D on the ground that D in w_4 was not made from ...[the] former parts [of a **table**].

This reply fails. If the juxtaposed situation were the result of simply **adding** features to the one in w_4, its description might have the implication for our description of w_4 [that] the essentialist attributes to it. But in specifying w_5 he **changed** certain details of my counterexample: in contrast to the smaller table in w_5 , D in w_4 has its origin in wood that composes a cleavage-product of block A. Characterizing this as a relevant difference between w_5 and w_4 is not tantamount to admitting that origin is essential. Rather, it is just acknowledging that the particular circumstances of a case matter for what we say about an object's identity. D's appearance and the conditions attending the start of its career in w_4 are sufficiently like B's appearance and the circumstances of its creation for us to say that D in w_4 is B; the similarities between the situation in w_5 and the above described one in w (or for that matter, the one in w_4) are not such that we would say that the second table in w_5 is B (or D in w_4). Surely, one can consistently hold that some, but not every, nonactual possible table whose coming into existence differs from B's is identical with B. (Price, pp.41-42.)

It seems that the origin essentialist will have to deny Price's intuition that the proportionately smaller table made from C would have been table B. Again it is not for me as an anti-essentialist to adjudicate between these conflicting intuitions. There is however one weakness in Price's argument against origin essentialism: while she makes a plausible case for the thesis that an object need not have originated from **all** of its actual original material, she has said nothing so far to suggest that an object could have originated from **little** or even **none** of its actual original material. At one place she dismisses the suggestion that an object must

have originated from at least a sizeable proportion of its original matter. Her ground for this dismissal can then be incorporated into an argument for the claim that an object could have originated from little or even none of its actual original material.

> Suppose that in the actual world w Brown uses nine equally heavy pieces of lumber to build table T. [Suppose the essentialist thesis] ... that T could not have been made out of the wood in fewer than six of the nine boards from which Brown actually makes T. .. Now, consider a non-actualized possible world in which Brown builds a table that is qualitatively indistinguishable from T by nailing nine boards together. Suppose that he draws on a pile of thirteen boards that contains the nine used to make T in w and that each of those nine boards is composed of the same wood that constitutes it in the actual world. Also suppose that by time t Brown has joined five of the nine from which he built T in w and three boards not among those nine. So, according to the essentialist ... [thesis] now under review, the addition of one more board may determine whether the finished table is T. Yet if the result of adding one of the boards that originally composed T in w is T, then intuitively, so would the result of adding one of the boards that were not actually used in the construction of T. In other words, whether the final product is T does not seem to hinge on the identity of the wood whose addition completes the table. Yet it would if the ... [essentialist] thesis about origin were true. (Price, pp.44-45.)

Price's main claim is that if the result of adding one of the boards that originally composed T in w is T, then intuitively, so would the result of adding one of the boards that were not actually used in the construction of T. So however low the threshold is set between a table's being T and its not being T, the table's identity cannot depend on the identity of the final part to be installed. Suppose a billion atoms compose the table. Assuming for the sake of simplicity that atoms really are literally atomic, i.e. indivisible, let us set the threshold at the lowest possible level, viz. one billionth. In this case the last part to be installed will also be the only part whose installation has any potential bearing on the identity of the table under construction. But by Price's lights it has no such bearing. Hence the table's identity does not depend on the identity of any part to be installed, and *a fortiori* it does not depend on the identity of its actual original parts.

To the line of thinking embodied in Price's main claim the origin-essentialist could

object that by parity of reasoning, since the addition of a hair from a grey pate to a bald pate, makes the bald pate the bearer of at least one grey hair, so does the addition of a hair from a pate which is not grey. In other words the reasoning emdodied in Price's main claim is invalid. Thus the Pricean argument that an object could have originated from little or none of its original matter fails.

3.9 FORBES ON THE ESSENTIALITY OF ORIGIN

To my mind the strongest case for the essentiality of origin is made by Graeme Forbes. I shall follow Penelope Mackie's reconstruction of Forbes' discussion in chapter 6 of his *The Metaphysics of Modality*. (see Mackie, P 1987, pp.183-188.) Suppose (1) that an acorn a_1 was actually planted at time t_1 in position p_1 and grew into an oak tree o_1. Suppose further (2) that o_1 could have grown from an acorn $\mathbf{a_2}$, where a_2 is distinct from a_1 and was planted at t_1 in p_1. Suppose again (3) that o_1 could have grown from a_1 but was planted at t_1 in $\mathbf{p_2}$, where p_2 is distinct from p_1. Suppose finally (4) that two oak trees o_2 and o_3 could have grown together, such that o_2's history was just like o_1's on supposition (2), and o_3's was just like o_1's on supposition (3). So o_2 would have grown from a_2, and o_3 from a_1. Since o_2 would have been distinct from o_3, both could not have been identical, if either could have been, with o_1. Now if o_2 would not have been o_1, or o_3 would not have been o_1, then this would have violated one or other of two principles about transworld identity and distinctness which Forbes endorses: viz. the principle of No Bare Identities and the principle that identity not be determined by extrinsic features.

According to the first principle, the difference between two qualitatively identical possible states of affairs cannot consist **solely** in the fact that they concern two distinct individuals a and b (Mackie, P 1987, p.173); i.e. if there are no differences in the identities of individuals other than a and b which ground the difference in question (Mackie, P 1987, p.181), then the two possible states of affairs cannot be

qualitatively identical after all, if, that is, they are indeed to concern **distinct** a and b. This of course is tantamount to an endorsement of PIOICP.

According to Forbes' second principle, intrinsic rather than extrinsic features should ground the facts of transworld identity and distinctness. An intended corollary is that it cannot be the case that a is identical with b just because nothing else is present to compete with a for identity with b (Mackie, P 1987, p.176).

To take o_2 first. If o_2 as described on supposition (4) would not have been o_1, then it would have been distinct from o_1 as described on supposition (2). But the only difference between supposition (2) and supposition (4) is that o_3 did not exist on supposition (2). Now the absence of o_3 would appear to be an extrinsic rather than an intrinsic difference, particularly since o_2 and o_3 are causally insulated from each other (see Mackie, P 1987, pp.185,191.) In that case the principle of No Extrinsic Determination would be violated. Alternatively it might be replied that no such violation has occurred; rather what grounds o_2's distinctness from o_1 as described on supposition (2) is simply a bare, i.e. irreducibly brute, difference between them, despite their qualitative identity and identity of career and material composition. But this would then be to deny Forbes' principle of No Bare Identities.

By parity of reasoning the denial of $\mathbf{o_3}$'s identity with o_1 leads to the same result, viz that one must deny Forbes' principles of transworld identity and distinctness.

So an affirmation of these two principles requires that it not be the case that the situations described on suppositions (2), (3) and (4) **all** be possible. The essentialist will typically disallow (2) rather than (3) or (4). In other words, while o_1 could have grown in a different position, and presumably in different soil, it could not have originated from an acorn other than its actual one. Its origin is essential to it.

Now as Penelope Mackie explains:

> All that one needs to do, ... , while accepting Forbes' ... principles about identity, is to assign **some** ... essential property to the tree. Nothing in the logic of the argument require that this property have anything whatsoever to do with the beginning of the tree's existence. Such a property might, for example, concern the exact location of the tree at some precise moment in 1950, long after its original planting. (Mackie, P 1987, p.186).

If this latter property rather than origin were essential, then the essentialist would probably urge that o_1 could have been as described on suppositions (2) and (3), provided of course that the respective descriptions included the circumstance that the tree had some precise location at some precise moment in 1950. Whichever of the trees as described on supposition (4) would not have met this description, could not have been o_1. But as P. Mackie observes:

> ...any such proposal would conflict with the intuition that it is not an essential property of any individual such as a tree that it **continue** in existence for any significant length of time. So ... [one is] driven to make use of features of the way those things originated. (Mackie, P 1987, p.186.)

While Mackie does not explain this, the reason why one is so driven is that any such proposal would conflict with a corollary of the above intuition. The corollary is that the object in question only **accidentally** has any temporally indexed property of being-F-at-t for any t and any F, such that t is any time t **after** the object's origination, and F is such that it characterises the object at t and 'F' involves no references to times **earlier** than t. This follows from the assumption that an object could have had a history shorter than its actual one. Moreover, since on this view, the conjunction of **all** of the properties of an object's origin constitutes the conjunction of all of the properties which are **essential** to the object, i.e. necessary for its transworld identity, it follows that the conjunction of all of these properties is **sufficient** for the transworld identity of the object; i.e. any object with all of these properties would have been the object under discussion. (In this way we have an argument for Noonan's proposal introduced in 3.7.3.1, the crucial assumption in Salmon's second argument for the essentiality of origin. But obviously this argument for Noonan's proposal already presupposes the essentiality of origin.) We

thus have an **individual essence** for the object, a specification uniquely essential to it.

It might be objected that since it is surely not an essential feature of an ordinary object that it came into existence in the first place, one cannot base such an object's essential properties on its origin. One would then have to construe its essential properties as **relational** properties involving reference to features of objects which **pre-exist** the object in question, its parents, say. But if, as the essentialist will most likely concede, its parents are not **necessary** existents either, then the parents do not provide the appropriate anchor for the object's essential properties. Nor presumably do the parents' parents provide the anchor, nor the parents' parents' parents, etc. But then we are led on a regress which presumably ends only at the very first assemblage of matter-energy at the beginning of the universe's history (a singularity). And if, as many essentialists would concede, the universe began to exist **contingently** and not necessarily, then the very first assemblage will not provide the appropriate anchor either, in which case we are left with no base on which to erect for any ordinary object **any** essential properties, let alone origin as an essential property. If, on the other hand, the essentialist insists that the universe began to exist **necessarily**, then the very first assemblage of matter-energy **will** provide the anchor in question. We would have found an ontological base for an ordinary object's essential properties, which properties are relational properties. But this then raises the perplexing question of how one is to determine the exact **nature** of the relevant relational properties. Which features of the first assemblage of matter-energy are to be related to the existence of a future individual in such a way as to secure the essentiality of that individual's origin? Of course if one denies the possibility of reference to future existents, then one cannot even construct the relational properties in question. (Thus recall Prior in section 3.5, who subscribes to a view of both time and facts as in passage; on this view the future is open and thus

any statement about objects-yet-to-be is ineligible for the assignment of a truth-value, since there is no determinate fact as yet to serve as a truth-maker for the statement in question.) And if the universe did not begin to exist but has an infinite past, then it is not clear how one even grounds the notion of an assemblage of matter-energy's having certain features essentially; perhaps an appeal to God becomes necessary.

The origin essentialist might reply to the preceding line of thinking that one can circumvent these difficulties by insisting that, while an object did not have to come into existence, it could not have come into existence with any origin other than its actual one. But clearly such a response would merely beg the whole question of the essentiality of origin. But the origin-essentialist isn't then forced to make the desperate move of declaring that, contrary to popular sentiment, ordinary objects are necessary existents. For the origin essentialist can simply point out that the reason why contingently existing objects have their origins necessarily is that this is secured by Forbes' twin principles of transworld identity. The question an opponent of origin essentialism must then address is whether one has good grounds to accept Forbes' twin principles of transworld identity, viz. the principles of No Bare Identities and of No Extrinsic Determination of Identity.

As for the latter principle, P. Mackie takes it that it is susceptible to two interpretations. On the first, a transworld difference in identities between x and y cannot depend on the fact that x is accompanied by a competitor for identity with y, or on the fact that y is accompanied by a competitor for identity with x. On the second, a transworld difference in identities between x and y cannot depend on the fact that x is accompanied by some individual or event causally isolated from y. (Mackie, P, p.192.) Mackie explains that the principle is very plausible on the second interpretation but not the first, because the **intraworld** counterpart of the

first interpretation runs foul of cases where identity over time legitimately depends on choosing between **competing** candidates, in terms of causal considerations. (Mackie, P 1987, p.192.)

What about the principle of No Bare Identities? Mackie concedes the preanalytic plausibility of this principle for someone who thinks that things could have been otherwise. Indeed she admits the appearance of bizarreness that would initially attend the rejection of this principle. (Mackie, P 1987, p.197.) And of course I have argued in chapter 2 that the essentialist ought to endorse this principle, which I there discussed as PIOICP. Yet Mackie ends up rejecting this principle presumably because she finds it initially much more compelling that (1) Fido, say, could not have had Rover's **actual** origin than that (2) Fido could not have had a merely **possible** origin for Rover. Her point seems to be that Forbes' argument for the essentiality of origin yields **both** (1) and (2), appealing as it does to the principles of No Bare Identities and No Extrinsic Determination. Mackie, who is not willing to reject the latter principle, and can accept (1) **without** (2) only, as she demonstrates, on pain of violating the No Extrinsic Determination principle, opts to reject the principle of No Bare Identities. (see Mackie, P 1987, pp.195-197.)

I think that Mackie is confused about the logical relationship of Forbes' twin principles of transworld identity to (1) and (2). What Forbes' argument secures with the aid of these principles and the further assumption that an object could have had a history shorter than its actual one, is that the only essential properties of an object are those of its origin. Therefore they are sufficient as well as necessary for its transworld identity. So any object with Rover's actual origin would have had to be Rover. Moreover, any object with any possible origin of Rover would have been Rover. This is (2), which Mackie finds objectionable. Presumably she finds it objectionable because she thinks that Rover has possible origins other than its actual

one, and that it is somehow metaphysically territorial that Rover exclude other objects from the possible possession of them. But given Forbes' supplemented argument Rover's possible origins collapse into his actual origin; there are no possible origins for him distinct from his actual one. Rover's actual origin is essential to his identity. This follows, to repeat, from the supplementation of Forbes' argument with the assumption that an object could have had a history shorter than its actual one. Without this addition the essentiality of Rover's origin (2) does not follow. So it is this assumption and not the principle of No Bare Identities that **Mackie** ought to fault, given her apparent desire to make accessible to dogs other than Rover the possession of possible origins for Rover. But now there is a further point. As long as some instantiated feature of Rover's actual history is made a sufficient condition of his transworld identity, then given the necessity of distinctness the possession of this instance becomes essential to his identity. Hence no other dogs possibly have this very instance, though of course they can have their own instantiations of the feature in question. So the real trouble, as Mackie should see it, is either the idea that there are sufficiency conditions of transworld identity or the necessity of distinctness. As it turns out she ends up rejecting the former. Indeed her discussion is originally premised on it. (Cf. Mackie, P 1987, p.176.) Of course a rejection of sufficiency conditions of transworld identity suggests the rejection of No Bare Identities. But it should be made clear that while Mackie's quarry is the principle of No Bare Identities, the logical culprits in the piece are the notion that there are sufficiency conditions of transworld identity and the necessity of distinctness.

Mackie even argues for the rejection of No Bare Identities from what she calls 'the popular conception of possible worlds as branching out from the past to the future':

> ... we are already committed to saying that there can be a possible world that is exactly like the actual world up to some point [t], after which the course of events goes differently from the way in which it actually goes. Yet if we can

> make sense of this, might we not also make sense of the idea of a possible world that is just like the actual world up until just before t, and then diverges from it in one respect,and one only, namely that it is Fido, rather than Rover, who comes into existence at t and has a career indiscernible from that of the actual Rover? (Mackie, P 1987, p.197.)

I would urge that the essentialist ought to respond by asking in virtue of what is it that the dog in question is Fido rather than Rover. Moreover, Mackie *ipse dixit*, confessing at an earlier place that

> it seems perverse to suggest that the world could have been **exactly** as it was in every respect right up to the moment before t, and that a dog could have been produced at t at the appropriate place, and yet that dog have been not Rover, but some other dog. (Mackie, P 1987, p.194.)

So given A(1) the principles of No Bare Identities and of No Extrinsic Determination of Identities; A(2) the claim that ordinary objects are necessary existents; A(3) the claim that an ordinary object could have had a history **shorter** than its actual history (excepting of course for the sake of consistency with A(2) objects which exist only for one instant); and A(4) the necessity of distinctness, it follows that an object's origin is both necessary and sufficient for its transworld identity. In other words, any object that could have had that origin would have been the object in question, and the object in question could not have had an origin other than its actual one.

As an anti-full-essentialist I reject the conclusion of this argument In particular, I reject A(3). The reason why I do so is that I do not accept that an object could have had a history shorter than its actual history. This is not because I believe that an object had to exist as long as it did, for I do not accept that either. Why? Because I find the notion of things' possibly being otherwise, which is presupposed in these 'could-have' and 'must-have' statements, unintelligible. And the reason why I find this notion and these statements unintelligible, is that in conjunction with

the premises of Chisholm's argument they entail the absurd, namely that one thing could have been another. (See chapter 2.) If they didn't entail the absurd, then Forbes would have a compelling argument for the essentiality of origin.

After all, if things could have been otherwise, then if I could have had an origin different from my actual one, say, different gametes, then in virtue of what it is the case that alternative gametes could have fused to form **my** zygote? Why would it have been my zygote, i.e. I and not something.or somebody else? Consider: if these alternative gametes had fused and the zygote had been allowed to develop, the resulting organism might have come to have a form, composition and a career (and personality even) identical to mine. Equally it might not have. But if it hadn't and we we still maintained that the resulting adult could have been I, in virtue of what would that adult have been I? Bemused silence in response to this question will tend to undermine or at least render puzzling the view that my zygote could have formed by fusion of alternative gametes, or that more generally an object could have had an origin other than its actual one.

In fact, the last question posed in the previous paragraph connects with a more general one which lies at the very heart of the issue of alternative possibilities for the one individual: viz. 'In virtue of what is it that the **same** individual is subject to different possibilities?' The answer by now should be obvious for an essentialist who endorses A(1) to A(4): it is the object's possession of its individual essence, viz its origin.

Kripke notoriously says that we simply stipulate whom we're talking about when considering an individual as a subject of alternative possibilites. But what ensures that the possibilities we entertain are genuine **metaphysical** as opposed to **conceptual** possibilities for that individual, the individual whom we've stipulated as

the object of our reference? What, in other words, makes ontologically and not merely stipulatively the individual in question **that** individual and not another?

Compare Sarker on this score, who represents the essentialist as urging that one could not assert that 'Russell could have been born of Mr and Mrs J.S. Mill without asking 'in virtue of **what** would the individual born be Russell?' Should a mere stipulation be enough or plausible?' (Sarker, p.141.)

Wiggins thinks that we do not have to find something in virtue of which the object of speculation is Russell. Wiggins does concede that the speculator has to be able to **rebut** the charge that they have lost the subject of discourse if they change its parents or origin. But:

> can ... [they] not rebut the charge by claiming to speculate about how [speaking now of Caesar instead of Russell] **the man whom Brutus murdered in 44 BC** would have fared if (say) Marius had been his father? To rebut the charge of losing the individual perhaps there must always be available to the speculator, **consistently with his speculation**, some such specification of whom he means by Julius Caesar. But ... this principle ... which is perhaps dubious ... does not favour any particular specification of who Julius Caesar is. (Wiggins 1974, p.335.)

Wiggins appeals to specifications that are thought to have actually held of Caesar, as a means of securing the identity of the subject of counterfactual speculation. He resembles Ayer in this respect, who proposes that we can consistently suppose anything whatsoever to be true of an object, provided that we retain some foothold in its actual history (Ayer 1975, pp.182-183):

> If we identify ... [the object] by its actual origin, we can think of it as occupying a different position at the present time; if we identify it as the ... [object] which is now in this position, we can think of it as having had a different origin (Ayer 1975, p.184).

The point, for Ayer, of seizing on a specification that the object actually meets, is not to underline its status as an essential property of the object, but to secure reference to the object as a subject of counterfactual speculation. But it follows from my discussion that Ayer would not endorse all of A(1) to A(4). For if he did, he would have to concede the existence of individual essences, in particular the object's origin. But this is what he denies.

So if things could have been otherwise, then A(1) to A(4) would provide a compelling case for the essentiality of origin. But, as I have been at pains to argue in chapter 2, it is not the case that things could have been otherwise, in the same way that it is not the case that green dreams sleep furiously. So Forbes' argument for the essentiality of origin fails. Nevertheless its superiority to other arguments for the essentiality of origin, such as Kripke's or Salmon's, is evinced in the consideration that where it fails, these other arguments also fail, in **addition** to failing elsewhere.

CHAPTER 4 THE ESSENTIALITY OF SORTAL SATISFACTION

4.1	INTRODUCTION	188
4.2	UNITY AT A TIME AND OVER TIME	189
4.3	SORTALS	191
4.4	SORTALS, NATURES AND NATURAL KINDS	199
4.5	SORTALS AND CAUSALITY	204
4.6	PUTNAM ON THE ESSENTIALITY OF SORTALS	210
4.7	BRODY ON THE ESSENTIALITY OF SORTALS	219
4.8	WIGGINS ON THE ESSENTIALITY OF SORTALS	222
4.9	THE OMNITEMPORALITY THESIS	228

CHAPTER 4 THE ESSENTIALITY OF SORTAL SATISFACTION

4.1 INTRODUCTION

An individual's origin is unique to that individual. So if it were essential to them, it would count as an individual essence of that individual. If by contrast being human were essential to Socrates, it would not count as an individual essence of him, for presumably Plato would share that essential property. 'Human' as a noun is of course a natural kind term, which in turn is a species of sortal term. Similarly, 'alligator' and 'horse' are natural kind terms and thus sortal terms. By contrast, 'car' and 'table' do not standardly count as natural kind terms, though they are sortal terms. I shall have more to say about natural kind terms and sortal terms later in this chapter. Now we recall from section 2.3.2 that Plantinga expresses agnosticism about the possibility of Socrates' having satisfied sortals other than those he actually did satisfy; in particular, he expresses agnosticism about whether Socrates could have been an alligator for example. Similarly, in the same section Cartwright was quoted as asking whether the horse Dancer's Image is necessarily a horse. In this final chapter I want to take up in greater depth the issues bearing on essentialism which emerge from a consideration of the role which sortals are said to play in our standard accounts of identity. In particular it is claimed that the sortals which an object falls under or satisfies are such that the object in question essentially or necessarily falls under or satisfies them. My overall aim will be to examine the case for the claims that, for example, Gauss had to be a man, or that Socrates had to be human, or that Dancer had to be a horse. I want to argue that there is no cogent case to be made for the thesis that an individual of which a sortal term is correctly predicable possesses essentially the property corresponding to that term. I shall argue that in the sense in which something which is sortally F **can** be granted to be

necessarily F, this involves at most the logical necessity of stipulative definition, and as such has no bearing on the further question whether something which is F is necessarily F in a **metaphysical** sense. Again, as I did for most of chapter 3, I will grant the essentialist, for the sake of argument, the coherency of the notion of alternative metaphysical possibilities, which notion underlies the essential/accidental distinction.

4.2 UNITY AT A TIME AND OVER TIME

As a point of entry into an account of the role which sortals are said to play in theories of identity, it will be useful to consider Hirsch's explanation, in his highly instructive *The Concept of Identity*, of just why it is that we need sortals in the first place. As my starting point I take the notion of a unitary material object, i.e. an aggregate of matter which constitutes a unit. There are two aspects of an object's unity: spatial and temporal. An object can be said to extend through space (**at** a time) and to endure **over** time (as it extends through space). Consider first an object's spatial unity. Thus for example the dog lying beside me **now**, as opposed to tomorrow or yesterday, chewing her plastic bone as I type, clearly counts as a spatial unit. But so could this dog *plus* plastic bone count as a spatial unit. In fact, any aggregate of matter, however discontinuously spread in space and time, counts on one understanding of 'unit' as a spatial unit. But this is not my understanding of 'unit' when I say that my **dog** now counts as a unit, but not my **dog *plus* her bone**. On my understanding, which I believe to coincide fairly closely with the pre-analytic notion of an ordinary object, a spatial unit, i.e. an **authentic** spatial unit, is characterised by the properties of what Hirsch calls 'spatial continuity' and 'dynamic cohesiveness'.

Hirsch describes the **spatial continuity** of an aggregate of matter x as the condition where any two parts of x can be connected by a continuous curve every point of which touches x. (Hirsch 1982, p.97.) But given an atomic theory of matter no macro-object will count as spatially continuous on this description, because every macro-object includes intermolecular interstices. Perhaps we could rectify Hirsch's formulation of the condition of spatial continuity by inserting a reference to the connectibility of any two parts of x **as presented phenomenologically**, in particular as presented in our visual and tactile fields. I will not pursue the matter, for it seems to me that what primarily marks off an aggregate of matter as a spatial unit is its capacity to **cohere over time**, the capacity for all its parts to remain **together** under various pressures. (cf. Hirsch 1982, p.237.) Thus a car tends to **move** (which takes **time**) **cohesively** and hence as a **unit**, unless it is subjected to a dispersive force. Or again a tree, while not generally separately movable, tends to **cohere** when subjected to various forces or pressures. (Hirsch 1982, p.108.) (This citerion of cohesiveness bears a striking similarity to Descartes' definition in *Principles of Philosophy*, part II sec.25, of 'one object' as that all of whose parts move together.)

Several points need to be made. Firstly, it is not clear that either spatial connectedness or dynamic cohesiveness is either necessary or sufficient for an object's spatial unity in all imaginable circumstances. Nonetheless these two concepts are, as Hirsch suggests, likely to appear in any general analysis of an object's unity through space. (Hirsch 1982, p.237.) Secondly, in explicating the concept of **cohesion** we have appealed to the rather **vague** notion of **togetherness** at each instant in an interval of time. Thus an aggregate of matter counts as an authentic spatial unit at a certain time, despite adjustments to prior deforming pressures, provided its parts are **together** enough; any parcel of matter that is not close enough fails to belong to this spatial unit, even if previously it was close

enough to belong. How, then, do we determine more precisely what is to be regarded as one aggregate, and what is to be regarded as a separate aggregate? The answer, as we shall presently see, is provided by an appeal to **sortals**. Thirdly, in thus appealing to cohesion **over an interval of time** the condition of dynamic cohesiveness presupposes an account of the conditions of the **temporal** unity of a material object. And here a vicious circularity in this account of spatial unity threatens to emerge; for standardly, and here Hirsch's account is no exception, the temporal unity of an object is accounted for in terms of certain relations which obtain between one **spatial** unit and another. For an aggregate of matter is said to cohere over an interval of time only in as much as at various (most) instances in that interval it remains more or less a spatial unit. Perhaps the circularity is unavoidable in that the notion of the spatial unity of an object at a time may prove to be unanalysable. In any case I do not need to solve this problem here, for my primary concern is not with an analysis of the concepts of unity at a time and identity over time as such, but rather with the claim that the sortals which an object satisfies are essential to its unity and identity. I think this latter issue can be pursued profitably without a rigorously developed account of the notions of unity and identity. In any case, considerations of space dictate that this will be my approach.

4.3 SORTALS

So given that the spatial unity of an object seems to be tied up with its temporal unity, let us turn to consider this latter notion. As Hirsch explains, the temporal unity of a standard object is constituted by spatio-temporal and qualitative continuity under a sortal. (Hirsch 1982, p.99.) That is to say that among the general terms of a language like English, there are some which earn the designation of 'sortal term' or 'sortal'. They do so on the basis of their capacity to mark off just

those authentic spatial units whose relations over time constitute temporal units. Thus 'Staffordshire bull terrier' counts as a sortal, because any aggregate of matter fitting this description at **one** time counts as an authentic spatial unit, and furthermore two such units existing at different times will count as numerically the same object - as being temporally unified so to speak - just where one is spatio-temporally continuous with the other. Hence since the pup I brought home at two months was a Staffordshire bull terrier, and is spatio-temporally continuous with the Staffordshire pup that is now eight months, the two month old pup is one and the same object as the eight month old pup. This holds notwithstanding the great disparity between them in material composition, size and appearance.

Hirsch explains that a sortal S is typically associated with the idea of a certain kind of size or shape or internal make-up which an object must exemplify in order to qualify as an S-thing. (Hirsch 1982, p.101.) Note that the only sense of 'must' in the previous sentence which I can grant is a logical 'must': being S is **defined** in such a manner that any object not satisfying that definition is disqualified by logic from being S. It is my contention that there is no cogent case to be made for the further claim that an object that is S - where henceforth all occurrences of 'S' (or 'S_1', 'S_2', etc.) in the remainder of this chapter refer to a sortal - **must** be S in some **metaphysical** sense. This is compatible with there being certain conditions that an object must meet **conceptually** or **logically** to count as an S.

To say that there are certain conditions which an object must meet even just conceptually as opposed to metaphysically in order to satisfy a sortal S, is to say that there must be explicitly available to users of natural languages a certain set of conditions which are severally necessary and jointly sufficient for an object's satisfying S. Schwartz describes this as a part of the traditional Lockean theory of

the meaning of sortal terms. (See Schwartz 1977, p.14ff.) He casts Putnam as an opponent of this theory. Yet Putnam does not reject the requirement that there be associated with the use of a natural kind sortal a set of necessary and sufficient conditions. It is just that he takes these conditions to involve theoretical properties rather than empirically manifest ones.

Putnam insists that the conditions of use associated with a sortal such as 'lemon' involve theoretical rather than phenomenological properties of lemons. Putnam insists on the theoretical rather than the phenomenological properties because of the problem of atypical specimens. For example, a lemon still counts as a lemon even if it lacks one of the following characteristics of lemons, which on the traditional semantics are thought to define the kind term 'yellow': being yellow in colour, having a tart taste, having a certain kind of peel, etc. After all, while these characteristics characterise typical members of the natural kind, the natural kind may have **abnormal** members. (Putnam 1970, p.103.) So while an object described as a lemon usually shares a sufficient number of properties with the normal or stereotypical members of the class of lemons, Putnam urges that none of these stereotypical properties need be possessed by an object to qualify as a lemon. This is because the stereotypical properties are empirically manifest properties, qualities like yellowness or tartness, and according to Putnam, empirically manifest properties do not feature in the definition of the kind term. What do figure in the definition, according to Putnam, are those **fundamental** properties of the kind which are causally responsible for these empirically manifest properties, among other things. As McGinn puts it, the fundamental properties of a kind underlie these other properties; the latter are nomically derivative from these fundamental properties, because they have their basis in them. The fundamental properties explain the **derivative** properties. (McGinn 1975, p.179.) Modern science of course identifies

these fundamental properties with those of the microstructure associated with the kind. This microstructure amounts in other words to the object's Lockean real essence. The important point to be highlighted is that Putnam has it that an object's microstructure as investigated and discovered by scientists is **definitionally** and so **conceptually** essential to the object's being of a certain natural kind. So an object completely unlike a lemon in appearance may nonetheless qualify as a lemon if it shares the microstructure definitive of the kind. Conversely, an object exactly like a lemon in appearance may yet fail to qualify as one if it lacks the microstructure definitive of the kind. So, since Putnam requires that there be necessary and sufficient conditions associated with the use of natural kind terms, he does not, *pace* Schwartz, diverge from the traditional Lockean semantics.

It is worth emphasising that on Putnam's view, natural kind terms and more generally sortal terms are not generally applicable on the basis of a set of necessary and sufficient conditions which make reference to an object's **derivative** properties. Recall that the derivative properties are those that depend causally on other properties of the object, the fundamental properties. On Putnam's theory, it is the **fundamental** properties that enter into the definition of kind terms. One can notice in passing that this observation might appear to hold less plausibly for **artifactual** kinds. For artifactual kinds are defined in terms of function, and function is a **derivative** property, for it supervenes on the spatial arrangement of the artifact's consituent parts. But since the same function is realisable in many different materials and configurations (think of the different ways of constructing a table), the properties on which function supervenes are **open-ended**. Consider for example a table: a table by definition supports articles for various purposes. But does this function require that a table have, say, at least one leg? For example, would a solid cube count as part of a **legless** table if a platform mounted on the cube jutted out far

enough for people to place their knees underneath? Or would the cube have to count as the one and only **leg** of the table? Or consider a clock: does it have to have springs, or hands, or electronic circuitry to perform its function of timekeeping? Sand will do just as well, as will the motion of the sun. But then no condition is by itself necessary. Instead we have a possibly open-ended disjunction of conditions such that no one condition is necessary even though it is sufficient for performing the function in question.

One can sympathise with the rationale of Putnam's account, for as far as **derivative** properties as opposed to fundamental properties are concerned, sortals do not as a rule, particularly not with respect to organisms, come complete with a closed set of conditions, which are jointly sufficient and severally necessary for the application of the sortal. (This point is borne out by consideration of the glaring vagueness of standard dictionary definitions of natural kind terms in terms of manifest properties, and the concomitant uncertainty about whether to classify as analytic or not various propositions involving those terms.) Instead one finds associated with the use of a sortal S certain so-called **normal** or **typical** or **characteristic** instances, **stereotypes**, from which S-things may diverge in certain respects to different degrees.

What I want to do for now is to address the question of what we mean when we say that something is a **normal** S-thing. This is a question about classification. One must have an account of classification is one is to have an analysis of the concept of a normal S-thing. Now it is highly likely that all macro-objects are as a matter of empirical fact qualitatively discernible; so one wonders on what basis it is that one object gets to be classified under the same head as another object. Certainly we are limited by what we can perceive, and in this respect the sensory and conceptual

equipment which evolution has selected plays an important role. But why have natural languages tended to label certain perceptual clusters and not others? Presumably because certain clusters have been noticed simply because they have tended not only to appear and then reappear with great frequency and/or regularity in everyday experience, but also to prove important in the actions and interactions of conscious agents. The reason that they regularly appear is that they evince the operation of causality. The world is so ordered that similar causes produce similar effects; given the existence of sufficient similarities in various respects, both at a time and over time, among distinct phenomena, similar effects will regularly follow similar causes. Nonetheless I shall argue in the last section of this chapter that there is a part to be played by subjective considerations in determining the perceptual clusters that correspond to sortals.

In any case, what I want to argue in this paragraph is that if we consider an object without reference to other objects, and think of it, moreover, just as the sum of its properties at **each** moment of its **entire** career, then the notion of that object's being normal or abnormal ***qua* that object** is of little point. For if an object is to be described as normal or abnormal, it must be **compared** with some standard. But then other objects could be compared with this same standard. One might even choose one or more of these objects as exemplifying the standard and then gauge the degree to which other objects resemble these exemplars. In any case, one will have in mind a certain cluster of features, usually at the phenomenological level in the first instance, in virtue of whose possession the exemplars are determined to be exemplars in the first place. For example, one cluster might comprise a range of colours in the yellow spectrum; a range of ellipsoid-like shapes; a range of certain acidic tastes; a range of certain sizes, etc. An object is then compared with these exemplars and then designated as normal or abnormal, according as it resembles or

diverges from them. When the object compared is determined to diverge too much, it is then disqualified from being an S-thing. Clearly, an S-exemplar would be determined to be normal because it resembles one of the exemplars, namely itself, exactly. Now it would be of little point to compare an object, say x, with itself and say it is normal relative to a cluster whose elements comprise every one of x's properties over its entire career. For it would then seem that by default any object other than x will be abnormal relative to this cluster. But the whole point of the normal/abnormal distinction is to provide for the possibility of more than one normal object, and this possibility would be precluded if the comparison cluster comprised **all** of x's properties over its **entire** career. One would then have general terms which only one object ever satisfied, so natural language would become impossibly unwieldy both in predication and in reference. Thus one wouldn't be able to say 'a cat ran across the road' in both the case of a Siamese and a domestic cat, rather one would have to invent different appellations for them, and different appellations again for different Siamese cats. The upshot then is that there is little point then in calling an object 'normal ' ***qua* that object.** What we typically want to know is whether an object is a normal S relative to a cluster which is much narrower than the set comprising all the properties of some object. Of course what we primarily want to know is whether an object is S or not. But to determine this we have to determine whether the object in question is normal or abnormal for an S. If it is fairly normal, then it is an S; if it is too abnormal, then it is not an S.

So what is normal for objects that satisfy some sortal S is represented by a cluster of paradigmatic features associated with S. These features typically include at the phenomenological level reference to shape, size, colour, function, etc. The more theoretical sortals of the sciences typically include reference to phyical, chemical and biological constitution, interaction and activity, etc. Sometimes the one sortal has

both an ordinary language interpretation and a scientific theoretical interpretation. Often in such cases some of the information associated with a theoretical interpretation of a sortal becomes incorporated into the information associated with the ordinary language interpretation. In any case since sortals pick out natural units which undergo changes over time that happen to be governed by causal laws, the cluster of properties definitive of a normal S-thing must be serial. That is to say that the cluster definitive of a normal S-thing will in fact comprise a series of clusters arranged in order of increasing posteriority, where each cluster in the series represents the range of normal S-things at a particular stage of its existence. Thus five feet might count as within the range of normal heights for an elephant of two months, but certainly not of two years. Of course it is probably only in the sciences that this series is actually explicitly analysed into a detailed set of stages. Most ordinary language sortals only make vague reference to broadly delimited stages in the career of the objects which instantiate them. (And corresponding to some of these stages there are often **phase** sortals, e.g. 'puppy', 'kitten', 'boy'.) The point is that sortals have associated with them an idea of what counts as a fairly normal evolution or existence. For example one does not expect normal tables to disappear suddenly every five months and reappear the next day. And while some sortals such as 'table' are uncommitted about the normal length of the existence of their instances, some are explicit about this; thus we say that a normal cat usually lives about thirteen years. (I am not suggesting however that 'has an average life-span of thirteen years' is **part** of the **meaning** of 'cat'.)

4.4 SORTALS, NATURES AND NATURAL KINDS

Now this broad account of sortals has been applied by some in defence of full-essentialism. To this full-essentialist end the notion of a **nature** is introduced. On this view, every object has a nature. Its nature is what it is about the object that accounts for its characteristic functioning or development, in other words, what **fundamentally** it is in the object which causes a normal object's behaviour. The nature of an object has been variously identified with its matter (Presocratics), its form (Aristotle), or the arrangement of its constituent parts (see Aune, p.349), as exemplified in Locke's conception of an object's nature, which he calls 'its real essence'. (I do not follow Armstrong in his rather deviant usage in *Universals An Opinionated Introduction* of 'nature' as meaning the set of an object's intrinsic (non-relational) properties.) Modern science has followed Locke in identifying the nature of an object with the fundamental properties of its microstructure. Now while Aristotle did not avail himself of the idiom of sortals but rather of substance or essence, for Aristotle the nature of an object is indicated by what is in effect its associated **sortal**. In keeping with this tradition, it is commonly said by philosophers to be (in) the **nature** of things which are sortally S to undergo the changes characteristic of **normal** S-things. (see Moravscik, p.589.) Other general terms which are true of an object that is sortally S but which are not **sortals** associated with that object, do **not** indicate its nature, or, as it is often put in Stagirite idiom, do not indicate the **what-it-is-to-be** of that object. So on this sortal view of natures, under normal circumstances things which are sortally S will display features associated with S- stereotypes; and in abnormal circumstances the features of things which are sortally S will diverge from these stereotypical features; in either case these features are accounted for by the **nature** of the object.

For a rationalist these natures can be known *a priori*; for an empiricist they can be known, if at all, only *a posteriori*. Locke's scepticism about the possibility of the knowledge of natures or real essences leads in his semantics to an emphasis on the nominal essence, i.e. the stereotypical manifest features associated with the use of sortals, as the basis of classification of objects. Since Locke is sceptical about claims about the character of an object's nature, he cannot rule out the possibility that two objects which are sortally S in virtue of exhibiting a range of features associated with certain S-stereotypes, may for all that have different natures or real essences which just happen to cause the same ranges of features. (See Ayers 1981, pp.255&257.) This undermines for him an Aristotelian view of kinds which holds that objects classified together in virtue of displaying the same nominal essence also share a real essence. Of course, modern science has not confirmed Locke's sceptical fears, for we now have good reason to believe that in the natural sciences at least, fundamental classes such as various kinds of subatomic units, of atomic units, of molecular units and some cellular units are composed of objects which share the same nature or real essence. (For Locke of course, fundamental **particles** or corpuscles would not have a real essence.)

Now this notion of a nature has been used as a criterion of the notion of a natural kind. A natural kind is minimally a class of objects which bear some **special** similarity to each other. (See Sober 1995, p.345.) Wilkerson tries to provide an account of this special similarity in terms of natures or real essences:

> If the notion of a natural kind is to be at all interesting, two conditions must be fulfilled. First, the notion of a natural kind must be tied to that of a real essence. That is, whether we are talking about kinds of stuff ... or kinds of individual ... , members of natural kinds have real essences, ... properties that make them members of the relevant kind, and without which they could not be members of the relevant kind. ... Second, members of natural kinds, and the corresponding real essences, lend themselves to scientific investigation. It is possible to have a science of gold ... or of tigers [or oaks]... , because it is possible to make suitable theoretical generalisations about their behaviour. It

> is precisely because gold has a certain ... [real essence] that it has certain properties (its being malleable, fusible, etc.); it is precisely because an oak has a certain genetic constitution that an oak has certain properties (a characteristic way of growing and reproducing itself) ... . In contrast it is impossible to have a science of ... [non-natural kinds. There] are no very definite constraints on their properties and behaviour. ... [N]atural kind predicates are inductively projectible, whereas other predicates are not. If I know that a lump of stuff is gold, or that the object in front of me is an oak, I am in a position to say what it is likely to do next, and what other things of the same kind are likely to do. I know for example that the gold cannot turn into water, and that the oak will not in due course produce tomatoes. ... Certain outcomes are ruled in, and others are ruled out, by the real essences of gold or oaks. In contrast, if I know that the stuff in front of me is rubbish, or that the object over there is a table, I am in no position to say what it is likely to do next, nor what other things of the same kind are likely to do. ... In short, because there are no real essences that make rubbish rubbish, and tables tables, I cannot even in principle make sound inductive projections about rubbish *as such* or tables *as such* [my emphasis]. ... I can make safe predictions about the behaviour of my table or rubbish heap ... [only by] exploiting the fact that every object, or quantity of stuff, will belong to at least one natural kind, even if it also belongs to one or more non-natural kinds. My predictions rest, not on any facts about membership of non-natural kinds, but on facts about membership of natural kinds. I am not predicting the behaviour of the table *as a table*, but as a quantity of cellulose, and I am not predicting the behaviour of the rubbish heap *as a rubbish heap*, but as a quantity of decomposing vegetation of various familiar species. (Wilkerson 1988, pp.29-30.)

Now Wilkerson is just simply wrong when he denies that tables or rubbish have a real essence. For modern science takes it that they do: in the case of a table it is the microstructural properties of a table which explain its other properties. Similarly in the case of a piece of rubbish. So Wilkerson must be appropriating a rather special sense of 'real esence' when he denies that tables have real essences. The textual evidence suggests that when Wilkerson says that an object belonging to the extension of a kind term K has a real essence *qua* being a K-object but not qua being a K# object, what he means is that **all** the members of the extension of K have the **same type of real essence**, but that not all members of the extension of K# have the same type of real essence. That is why he allows an elm to have a real essence *qua* an elm, but not *qua* a tree.

With this Wilkersonian understanding of a real essence, we note that Wilkerson goes

on to argue that the distinction between natural kinds and non-natural kinds does not coincide with that between natural kinds and conventional kinds or between natural kinds and artificial kinds. For on the one hand there are many natural kinds whose members are produced artificially, e.g. kerosene and polystyrene, and on the other hand there are many non-conventional and non- artificial kinds which fail to qualify as **natural** kinds because their members neither have a real essence *qua* members of that kind nor do they lend themselves to scientific investigation *qua* members of that kind. Examples include: trees, shrubs, perennials, cliffs, mountains, clouds and thunderstorms. As Wilkerson observes, trees as such do not have Wilkersonian real essences, but elms, oaks and beeches do. Similarly, cliffs as such do not have Wilkersonian real essences, but sandstone, limestone and granite do. And clouds as such do not have Wilkersonian real essences, but gases, especially water vapour, do. (Wilkerson 1988, pp.31-33.) Dupre 1981 develops this point with respect to many kind sortals recognised by common sense but not by science.

Wilkerson concludes illuminatingly that

> the word 'natural' is misleading. As we have seen, trees, shrubs, annuals and perennials are all natural in some perfectly ordinary sense of 'natural', so are pebbles, volcanoes, rivers, glaciers, cliffs, thunderstorms, anti-cyclones, fronts. But none has a [Wilkersonian] real essence, and none is likely to yield any relevant theoretical generalizations. A perfectly ordinary sense of 'natural' is not enough [to secure the distinction between natural and non-natural kinds.] (Wilkerson 1988, p.34.)

When it comes to delimiting natural kinds, an appeal to the Wilkersonian criterion of a real essence - viz. having the same type of real essence (in the standard Lockean sense of 'real essence') - helps explain the biologist's practice of grouping together as members of the one species, and thus satisfiers of the same sortal, those organisms which can interbreed. After all, in some respects a mature and mentally functional member of *homo sapiens* is behaviourally more similar to a chimpanzee

than to a mentally defective human, notwithstanding the inability of the human and the chimp to interbreed. The criterion of having the same type of real essence isolates genetic structure, which as it turns out can be transmitted only by intercourse between organisms with sufficiently similar real essences.

So we can grant that the respective microstructures of objects said to satisfy the one natural kind sortal may be very similar; indeed, in the case of genetically identical organisms they are likely to be even more alike than usual. Nevertheless, I want to take issue with a suggestion implicit in Wilkerson's discussion that natural kinds must be non-conventional. For I believe that there exists a certain conventional element at least in the **boundaries** we draw which separate one natural kind from another, at any rate as far as organisms are concerned. After all, organisms that can interbreed typically do not have **identical** genetic structures and thus **identical real essences** but only **similar** genetic structures and thus only **similar real essences**. And as Odegard points out, the existence of continuous gradations of differences between individual organisms leads Locke to reject a facile identification of natural kinds with classes whose members share a real essence. For Locke rejects the Aristotelian view of fixed and immutable kinds. Locke does **not** think that two specimens of the same species cannot differ as much as two specimens of different species; he does **not** think that different species are so distinct that cross-breeding is impossible; he does **not** think that no two species are such that there can be an intervening hybrid species or monsters or changelings; he does **not** think that no individual can be a member of more than one species. (See Odegard 1975, p.208.) Thus while terriers and greyhounds undeniably have many features in common, the classification under 'dog' of terriers and greyhounds seems to reflect in part human decision. Or turning to chemical kinds: consider the chemist's practice of classifying under the one head of 'water' both ordinary water and heavy water. These all seem

to constitute examples of a certain relativisation of classification to human interests or decision and thus a certain conventionalism in our delimitation of natural kinds. I will develop this point with respect to sortals more generally in the last section of this chapter.

4.5 SORTALS AND CAUSALITY

Of course from the point of view of providing an account of the spatial and temporal unity of an object, the criterion of interbreeding is well-founded as a basis for determining natural kind sortals as far as organisms are concerned. For recall that on the standard view, an aggregate of matter is regarded as spatially unified when it tends to cohere over time, i.e. to remain more or less spatially unified, despite being subjected to various pressures. Now the reason why aggregates of matter generally cohere over time is that their behavior is governed by causal laws. These causal laws explain how the aggregate will behave in various circumstances; thus Newton's first law says in effect that a mass will change its velocity because of the causal influence of an unbalanced force. A set of biochemical laws might run to the effect that a system of plant cells will undergo certain chemical reactions if subjected to water, light and carbon dioxide. As a result of these reactions some new matter is incorporated and some old expelled. The expelled matter no longer counts as part of the spatial unit that is the cellular system or organism. In any case, the process is governed by causal laws. Indeed, it is just because the new matter becomes not just subject to but reacts on the causal processes of the organism that it now counts as part of it. And it is because the expelled matter becomes largely causally independent of the organism that it no longer counts as part of it. In effect, sortals pick out the different ways in which aggregates of matter can cohere over time through causal processes. It then becomes clear why the criterion of

interbreeding should play a significant role in determining which are the natural kind sortals we recognise. For reproduction is just another example of the kind of (complex) behaviour governed by causal laws which certain spatial units manifest.

In short, we tend to recognise an aggregate of matter at a time as an authentic spatial unit just where we expect it to cohere **over** time because of certain causal interactions belonging to it. These interactions explain its cohesion over time, its having an inside and an outside so to speak, and mark it off from its environment, its outside. The fact that interaction inside is constantly affected by interaction with the outside leads on the part of the object to a myriad of possible manifestations of law-governed behaviour, depending on the particular circumstances. This explains why S-things can behave abnormally, and thus why sortals do not as a rule have associated with their use a set of explicitly formulable necessary and sufficient conditions referring to **manifest** properties. These manifest properties tend to be **derivative** properties, whose manifestation is often a function of the interaction of the real essence of the object in question with the real essences of other objects. The best that one can say is that, for example, a lemon **tends** to display a certain range of colours, **tends** to display a certain acidity of taste and composition, **tends** to have certain shapes, sizes and textures at different times in its career. But what Putnam overlooks is that the same can also be said about the **real essence** of lemons. For individual lemons do not generally have **identical** genetic structures but tend to display a certain **range** of genetic structures. Hybridise too much and you no longer have a lemon but, say, a citrus of the lemonade variety.

The notion of causality that accounts for an object's unity at a time and over time explains two further features of sortalhood: namely, firstly, why some general terms such as 'tree' or 'car' or 'cow' qualify as sortals but others such as 'round object' or

'smooth thing' or 'red patch' or 'tree *plus* bird' do not; and secondly, why standard accounts of identity over time require spatio-temporal and qualitative continuity.

As to the second point, Hirsch takes it that our ordinary account of the identity over time of an object requires that changes in spatio-temporal neighbourhoods be **more** rather than less qualitatively continuous. That they cannot be **completely** qualitatively continuous is evinced by the consideration that when a tree loses a large bough in a storm it undergoes a qualitatively discontinuous - rather than a qualitatively continuous - change in shape; after all, the tree does not undergo increasing degrees of boughlessness until it is finally completely as opposed to only partially boughless. (See Hirsch 1982, pp.10-15.) Nonetheless we generally aim for more rather than less qualitative continuity in spatio-temporal **neighbourhoods**; of course, over **extended** spatio-temporal intervals identity is compatible with qualitative **dis**continuity (the baby is the same human being as the adult; the caterpillar is the same insect as the butterfly). So if for example the object in my visual field at midday is a VW, I rub my eyes for a second, and the object in my visual field is now a Porsche, I take it that the respective objects are not identical; I infer that the VW has moved on. What Hirsch does not explain however is **why** we tend to base our judgements of identity over time on the presupposition of qualitative continuity in spatio-temporal neighbourhoods. The explanation surely is that as a rule, objects do not undergo radically discontinuous qualitative changes in spatio-temporal neighbourhoods. And the explanation for this is surely that matter doesn't behave in that way, the causal activities constituting an object's nature do not as a rule produce radically discontinuous changes in an object's empirically manifest properties in **local** spatio-temporal neighbourhoods. Hirsch precludes the possibility of such an explanation given his view that causality is not fundamental to the notion of identity over time or even at a time. Yet I think that the notion of

causality **is** fundamental in this regard.

I turn now to the other point raised in the last sentence of the paragraph preceding the previous one, the point about why not all general terms qualify as sortals. Why, for example, does 'cow' qualify as a sortal, whereas 'red and white patch' does not? Again I think the answer lies in an appeal to causality. We normally want to delimit units at a time and trace their careers over time, but not just any old unit. The units that interest us in our non-philosophical moments are those we take to be authentic as opposed to contrived. And I take it that our criterion or at least our explanation of authentic unity just is the notion of independent causal activity. A cow for example causally interacts with its environment in such a way as to count as independent, at any rate independent enough to merit the description of being a substance in one Aristotelian sense, namely being ontologically independent. It is the nature of the object which we call 'a cow' which preserves the ontological integrity of the object for the duration of the object's existence. An object that is properly called 'a cow' endures moreover in virtue of its being a cow as opposed to being, say, a certain variegated patch. Should a red and white patch maintain some ontological unity over time, this will be not **because** of its being a red and white patch. As a red and white patch it owes its endurance to something other than its being **merely** a red and white patch; perhaps it owes its endurance to its being a red and white patch on the surface of a **cow**. It is the nature of the cow that explains the endurance of the red and white patch. The red and white patch is not ontologically because not causally independent of the cow. So I venture the conjecture that sortals are those general terms which when correctly predicated of an object point **directly** as opposed to indirectly to the causal activity or nature which constitutes the existence and persistence of the object. 'Cow' points directly to the causal activity constituting the existence of an object that is correctly

describable as a cow; 'red and white patch' by contrast points only **indirectly** to the causal activity constituting the existence of an object that is correctly describable as a red and white patch. Wilkerson develops a similar suggestion with regard to **natural kind** sortals: he takes up Aristotle's criterion of independent existence, and makes it the basis for the distinction between natural kinds and non-natural kinds, between those things that are what they are independently of other things, and those things that are what they are in virtue of a relation to something else. (See Wilkerson 1988, p.37ff.)

The notion of a nature as the internal cause of an object's behaviour is very much an Aristotelian notion (see Aune, p.349), as is the notion that sortal terms indicate these natures. For while Aristotle did not put the matter in terms of **sortals** but instead in terms of **real definitions**, real definitions are supposed to tell us what it is to be such a thing as the object in question (*to ti ein einai*), i.e what the object is **really** like (*kath hauto* as opposed to *kath sumbebekos*). (Aristotelian essences are of course general, not individual.) Moreover, real definitions express the object's *infima species*, which are just **sortals.**

Aristotle's own '*ousia*' , which he used to refer to an object's nature or real essence, has been rendered historically in some contexts by the latinised 'essence', in others by 'substance'. Now as Teller observes,

> Aristotle's notion of an essential property was that of a property that serves to give a unifying causal and scientific explanation of an entity's other properties. At the same time, Aristotle also clearly took an essential property to be a property such that nothing could lose it and yet remain the very same, numerically identical thing. (Teller 1975, p.248.)

In other words, there are two functions that historically the properties belonging to an object's nature are said to fulfil: firstly, they explain causally or in the case of geometric objects, logically, the object's **other** properties; secondly, they are

necessary to the identity of the object. So historically, there have sprung up two different notions corresponding to these two functions, but which confusingly have borne the same name, namely 'essence'. In the first sense of 'essence' corresponding to the first function, an essence is the object's nature. In the second sense corresponding to the second function, an essence is those properties which are such that an object is without them on pain of non-existence.

On the face of it it is not clear why an essence in the first sense (real essence) must coincide with an essence in the second sense (metaphysical essence). Are there any considerations which might ground the claim that an object with a certain real essence could not exist without it?

Now an object's real essence is sometimes described as **what** it is to be the object. And any one of an object's properties such that without it the object cannot be **what** it is - be the object that it is - is necessary to that object. But in a trivial sense of 'what it is to be an object', **every** one of an object's properties is such that without it the object would not be what it is. So it is important to bear in mind that what the sortal essentialist has in mind when they talk about **what** an object is, is not the **sum** of its properties at any time but a subset of those, viz. its **real essence**. And an object's real essence is indicated by the object's **sortal.** ('Cow' **directly** indicates an object's sortal even though it does not specify it **explicitly**; that is left to the zoologist.) So in asking whether an object's real essence must coincide with its metaphysical essence, we are in effect asking whether an object that is sortally S is necessarily S.

The idea is that an S object may undergo certain changes and survive them; but that should it suffer a loss of all or most of the properties constituting its real essence then it ceases to exist. So while a tree can survive autumnal defoliation it cannot survive being felled and cut up for timber, for it cannot survive the cessation of the metabolic activity of **all** of its cellular components.

4.6 PUTNAM ON THE ESSENTIALITY OF SORTALS

I want now to assess the essentialist implications of Putnam's position on the semantics of natural kind terms. Since Putnam wants to extend his position to artefactual kinds as well as natural kind terms, I am in effect going to assess Putnam's position on **sortal** terms in general.

We recall that the existence of abnormal members of a natural kind leads Putnam to reject the view that a natural kind term is definable in terms of a conjunction of manifest properties. Now as Ayer explains, the obvious answer to this objection is that

> what the list of characteristics defines is a 'normal member' of the kind in question. If a particular object fails to possess one or more of these characteristics it may still be reckoned a member of the kind, so long as it possesses the others. How far an object may deviate from the norm without forfeiting its membership is a matter for decision. It would seem too that not all properties can be accorded equal weight ...
>
> Putnam rejects this rejoinder. He argues first that 'the normal members of the natural kind in question may not really be the ones we *think* are normal' ... and secondly that 'the characteristics of the natural kind may change with time, possibly due to a change in the conditions without the "[real] essence" changing so much that we want to stop using the same word'...
>
> ... It is assumed that the meaning of a word which applies to members of a natural kind is uniquely associated with their '[real] essence', and this allows both for the possibility that at a given time the specimens of the kind that we regard as normal form an untypical sub-class of all those that have this essence, and for the possibility that the manifest properties of the genuinely typical members of the class change in the course of time. The meaning of the word stays constant in either case because of its constant attachment to the essence. (Ayer 1982, p.268.)

But how would we come to discover after time t that the specimen objects which we had hitherto taken to be representative of the kind K are in fact untypical? Presumably, this would be the case if (a) objects first examined after t proved to

share the same real essence as K-objects first examined before t but to manifest different derivative properties, and also if (b) there proved overall to be more of this later kind of K-objects than the earlier kind. But this presupposes that the same microstructure can ground different sets of derivative properties, and it is not clear how this can be the case - after all, same cause, same effect. Unless, that is, in point of fact we really have two microstructures each with their proper set of derivative properties. Perhaps the similarities in microstructure are close enough for specimens possessing either microstructure to merit the same generic designation. Heavy water and ordinary water are a case in point.

Similarly, the reason why later specimens display different derivative properties than earlier specimens would surely not be because the same microstructure has produced first one set of properties and later another, for again, same cause, same effect. The reason rather must lie in the fact that specimens in later generations have a different microstructure from specimens in earlier generations, but that for all that the respective microstructures are similar enough to warrant for various reasons the attribution of the same generic designation to specimens with either microstructure. A case in point is the progeny produced by sexual reproduction. Thus despite genetic differences from our parents we are still called 'human beings'.

Now as Forbes points out, Putnam's interest in natural kind terms is primarily in their semantics. (Forbes 1985, p.191.) Nonetheless Putnam makes certain claims which have been interpreted in essentialist terms. In particular he has been interpreted as saying that once we have scientifically discovered that the specimens of a kind K have a certain real essence E, then any individual that is a specimen of K has E essentially; the specimen has E necessarily in a metaphysical sense of 'necessarily'. But as Forbes explains, Putnam's claims concern natural kinds rather

than individuals, in particular with substances such as elements and compounds, and with biological species. Thus he is concerned with the natural kind *gold* rather than with individual nuggets or bars of gold; the natural kind *water* rather than puddles or glasses or lakes of water; the natural kind *tiger* rather than individual tigers in the zoo or jungle. It is to these natural kinds rather than their instances that Putnam seeks to attribute essential properties in the first instance. So according to Putnam, it is essential to water to be H_2O, it is essential to gold to have the atomic number 79, it is essential to the tiger that it be a mammal. (Forbes 1985, p.192.) Nevertheless, essentialists have found support in these sorts of claims for full-essentialism about the individuals which instantiate the kinds in question. Thus it is thought to follow that an individual nugget of gold could not have had an atomic number other than 79; a sample of pure water could not have been composed of anything other than hydrogen and oxygen; a tiger could not have been a fish.

There are two separate issues here. Firstly, does Putnam provide cogent grounds for the thesis that natural kinds have certain properties essentially? Secondly, even if he does, does this secure the corresponding full-essentialist thesis for individual instances of these kinds? In other words, does it follow that any specimen belonging to kind K, is such that it necessarily has the real essence E definitive of the kind term?

Let us turn to the first question first. Putnam says:

> Once we have discovered that water (in the actual world) is H_2O, *nothing counts as a possible world in which water isn't H_2O.* In particular, if a 'logically possible' statement is one that holds in some 'logically possible world', *it isn't logically possible that water isn't H_2O.* (Putnam 1975, p.233.)

The idea is that the molecular structure of a substance, or genetic structure in the case of biological species, is essential to the kind in question. And as Forbes

observes, should specific scientific theory prove to be wrong, we can replace references to molecular and genetic structures with references to fundamental natures, i..e those properties responsible for the other properties of the kind. (Forbes 1985, p.193.) Modern science happens to identify these fundamental natures with microstuctures. But if science is right, then whatever its merits with regard to natural kinds, Putnam's account shows little promise of applying satisfactorily to **artifactual** kind terms. For since artifactual kind terms are defined in terms of function, and the same function is multiply realisable, there will be no **one microstructure** necessary to explain the realisation of the function in question. Moreover, as Dupre has argued, the kinds recognised by commonsense in ordinary language often fail to correspond to kinds which science considers unified by a common real essence, i.e. to natural kinds by the Wilkersonian criterion. But as Forbes notes, this is not fatal to an essentialism of real essence for natural kinds, for the essentialist could still mount their case with reference to the species recognised by science. (Forbes 1985, p.194.) So what are we to make of Putnam's thesis of the necessity of microstructure or real essence to its respective natural kind?

Putnam argues for this thesis by way of his famous Twin Earth thought experiment. Explorers from earth investigate another planet apparently just like the earth. They discover a liquid substance A which while identical to the water on earth in respect of appearance, modes of occurrence and dispositional properties, proves to have a molecular structure of XYZ rather than H_2O. Putnam claims that any identification of A with the kind water is mistaken, even if the microstructures of both water on earth and the substance A on twin earth are discovered only after the discovery of the existence of A on Twin Earth. (See Putnam 1975, pp.223-235.)

Ayer analyses Putnam's argument for the necessity of microstructure thesis as follows. (Again, we take water as our natural kind without loss of generality.)

> (1) x is to be counted the same liquid as y if and only if specimens of x have the same microstructure as quantities of stuff which are ... [ostensively or otherwise identified] as specimens of y.
>
> (2) If y is water its specimens have the microstructure H_2O.
>
> (3) What is and always has been meant by the English word 'water', and its counterparts in other languages, is anything that has this microstructure, whatever its manifest properties may be.
>
> (4) Nothing that has a different microstructure, whatever its manifest properties may be, is such that the word 'water', or its counterparts in other languages, is properly applicable to it. (Ayer 1982, p.270.)

Ayer responds that (1) is prescriptive while (2), (3) and (4) are descriptive. Of the latter three he finds only (2) true. Since he finds (3) and (4) false, he finds little to recommend prescription (1). Ayer argues against (3) and (4) as follows.

> Suppose that in some part of this world we came upon stuff which had the chemical composition H_2O but did not have the properties of falling as rain, allaying thirst, quenching fire and so forth, perhaps even failed to appear in liquid form. I certainly should not call it 'water' and should be surprised if the majority of English speakers did so either. Conversely, I believe that most English speakers would still apply the term 'water' to stuff that had such manifest properties as I have listed, even if it had a different chemical composition. And the same would apply *mutatis mutandis* to speakers of other natural languages.
> ...we are presented with a set of properties which are habitually found in combination and ... so far as we know the stuff in which they are combined invariably has the chemical composition H_2O. It is not necessary, in any sense which I can understand, that these properties should occur together, or that all the specimens of the stuff in which they are combined should have the same chemical composition, or that this composition should satisfy the formula H_2O rather than some other. It is enough for any reasonable purpose that these generalisations should actually hold, and I judge there to be more loss than profit in any attendant talk of essence or necessity or possible worlds. (Ayer 1982, p.270.)

In effect Ayer is opposing his linguistic intuitions to Putnam's. But even if we accepted Putnam's prescription (1) and withheld the appellation 'water' from the liquid on Twin Earth, this as such has no deep metaphysical significance. It would merely reflect the logical consequences of a decision to adopt a convention, and conventions are neither true nor false; they correspond to no metaphysical fact of the matter. What would still remain unanswered is the question whether the kind *water* could have been, in some metaphysical sense of 'could', other than H_2O. After all, even if it is true that we would have described the contents of a tin of olive oil as water had the oil counterfactually been water and not oil, still it is a further question whether water itself **had** to be H_2O. It would follow logically from the linguistic convention, that any **actually** existing substance which is not correctly describable as water is not correctly describable as H_2O either. So trivially any such substance is such that it is essentially not H_2O in a **conceptual** sense. But it is altogether a further question whether the said substance is essentially not H_2O in some **metaphysical** sense. Essentialists often give the impression that a kind's being essentially F in a conceptual sense is sufficient for its being F in a metaphysical sense. But this is fallacious, perhaps due to a conflation of metaphysical with conceptual necessity.

For his part, Putnam thinks that his Twin Earth thought experiment does provide an answer to the question whether a kind that is essentially F in a conceptual sense is essentially F in a metaphysical sense. He says, continuing his thought experrment:

> Suppose ... that I discover the microstructure of water ... [to be] H_2O. At this point I will be able to say that the stuff on Twin Earth that I earlier *mistook* for water isn't really water. In the same way, if you describe not another planet in the actual universe, but another possible universe in which there is stuff with the chemical formula XYZ which passes the 'operational test' for *water*, we shall have to say that that stuff isn't water but merely XYZ. You will not have described a possible world in which 'water is XYZ', but merely a possible world in which there are lakes of XYZ, people drink XYZ (and not water), or whatever. In fact, once we have discovered the nature of water,

> nothing counts as a possible world in which water doesn't have that nature. (Putnam 1975, p.70.)

Now as Forbes observes, Putnam just seems to be assuming that if we grant that water is whatever in the universe bears a certain equivalence relation (viz. namely 'same liquid as')to the stuff actually identified on earth as water, then it follows that water is essentially H_2O. But that this does follow is hardly unproblematic. For even if the application of a natural kind term like 'water' within a possible world is determined by the fundamental physical properties of paradigm instances of the substance to be found at the place where the use of the term originated, this may still be compatible with the existence of a possible world where every sample of water has molecular structure XYZ. (Forbes 1981, pp.34-35.) At least Putnam has given no reason to think otherwise. After all, possible worlds are not just more planets like twin earth, only further away; they are causally and spatio-temporally isolated from the actual universe. So the relationship between Twin Earth and the earth is not clearly analogous enough to that between a possible world containing both the earth and Twin Earth and a possible world containing just the earth as it actually is, at least not clearly analogous enough to support an unproblematic extension of results in the one case to the other case. (Cf. Forbes 1985, p.194.)

Forbes considers whether an appeal to considerations adduced in his (Forbes') argument for origin essentialism might also secure the thesis of the essentiality to a kind of its microstructure. (Forbes 1985, p.195.) These considerations are of course Forbes' twin principles of identity: the No Bare Identities Principle and the No Extrinsic Determination iof Identity Principle, whose role in Forbes' argument for origin essentialism we examined in chapter 3. (See glossary.) Would the denial of the essentiality to a kind of its microstructure lead to bare tuths about the identity of kinds or to extrinsic determinations of the identity of kinds? In an earlier paper

Forbes considers with some agnosticism the suggestion that the denial of the essentiality to a kind of its microstructure leads to extrinsic determinations of the identity of the kind. He begins by asking what kinds of objects natural kinds are in the first place. What for example does the substance term 'water' designate?

> Not the set, or mereological sum, of all quantities of actual water, since, presumably, there could have been more or less water than there actually is. The only plausible candidate seems to be *waterhood*, the abstract property of being water ... [So] if the transworld identity conditions of an abstract object such as the property of being water are not extrinsically determinable, there must be something which fixes the identity of that property in a world ... And what is more natural than the suggestion that the identity of a substance property picked out with respect to a world by some of its characteristics there (e.g. that things which have it are thirst-quenching) is fixed by the identities of the fundamental physical properties of the things which have the neutrally described property? ... However, when we are concerned with abstract objects such as substance properties, there seem to me to be grounds for scepticism that ... [reasons to believe Putnam's essentiality thesis] will be forthcoming from an application of the intrinsicness constraint, to do with the question of whether the concept of transworld identity for such abstract objects is sufficiently well articulated to entail that someone who denies .. [Putnam's thesis] is committed to *falsehoods*. It may not be definitely *correct* to say that if matter had been non-atomic, so that there were no molecules anyway, but some substance had exhibited all the characteristics of water compatible with that, then that substance would have been water. But it may not be incorrect either; there may simply be no fact of the matter. (Forbes 1981, pp.36-37.)

Forbes has overlooked the possibility of the existence of bare truths about the transworld identity of kinds. The problem is this. What it would be like for properties to be otherwise than they actually are? If for example we take it that there is the property of being water, could it have been different and yet retained its identity as the property of being water? Suppose that there is a property of containing two hydrogen atoms and one oxygen atom Suppose furthermore that the property of being water consists in part of this property. Could it then have consisted in part of the property of having two oxygen atoms and one hydrogen atom, such that hydrogen atoms would have behaved just as oxygen atoms actually do and *vice versa*? Or is the supposition absurd? If the supposition is rejected

because it presupposes bare, i.e ungrounded, facts about the transworld identity and distinctness of the properties of being hydrogen and oxygen, then I take it that no properties could have been different. (Note how this line of reasoning parallels Chisholm's general argument against essentialism.) Hence the property of being water, which is the kind *water*, could not have been different. In particular it could not have had the property of having a molecular structure other than its actual one.

But this essentialist concession to Putnam on the issue of kinds, where kinds are being understood as properties, provides no resolution for the corresponding issue about **instances** of these kinds, namely the issue whether an object that is K has the rel;evant real essence essentially. This is so for two reasons. Firstly, I have argued that the great lesson of Chisholm's argument is that the threat of bare truths of identity and distictness means that not only could properties not have been otherwise, but that their instances also could not have been otherwise; that, indeed, the notion of a thing's being otherwise is incoherent. Secondly, even if we ignore this lesson and concede Putnam's thesis with respect to kinds, this by itself provides no reason to believe that an individual that belongs to a natural kind K has essentially the real essence E associated with K. For clearly, while it is true that an instance of K has E essentially in a **conceptual** sense of 'essentially', just where K itself has E essentially in a conceptual sense of 'essentially', it would be fallacious to reason that therefore the said instance has E essentially in a **metaphysical** sense of 'essentially'. We cannot draw this conclusion unless we accept **both** that the **kind** K has E essentially in a **metaphysical** sense, **and** that the instance in question is essentially K in a **metaphysical** sense. And on the question of whether the said instance is essentially K in a metaphysical sense Putnam's Twin Earth considerations have no bearing.

4.7 BRODY ON THE ESSENTIALITY OF SORTALS

I therefore turn to an argument by Brody which contends that an individual that is sortally K is essentially K. In Brody 1980 chapters 4 and 5, Brody adopts a model of counterfactual possibilities for actual individuals, according to which these possibilities, to qualify as genuine possibilities, must share at least some segment of the individual's actual career. As Penelope Mackie explains (Mackie P 1994, pp.314-315), on Brody's model something is a complete possible history for, say, Socrates only if it incorporates some segment, however small, of his actual history. (I shall follow Mackie (1994, p.317) in calling this 'the overlap requirement'.) The overlapping segment need not include Socrates' original state, for to do so would be to make **every** feature of Socrates' origination - including place of conception - essential to him, which, as Brody observes, many essentialists would find counterintuitive. (Brody 1980, pp.119-120.) The point of course for Brody in making the overlap requirement is that he tries to avoid the problem of the identity conditions of individuals across possible worlds. He does so by anchoring their identity, much in the way Ayer is quoted on page 188 above as suggesting, in segments of an individual's actual complete history. But as we shall presently see, this generates unwelcome results for the essentialist.

Now sortals are commonly held by the essentialist to be such that any instance of a sortal S is such that if it instances S at any moment of its career then it instances S at **every** moment of that same career. As Mackie puts it, once an S, always an S. (Mackie P 1994, p.314.) I will examine this assumption - call it 'the omnitemporality thesis' - in due course. The point is, as Mackie explains, that given this thesis and the overlap requirement, any possible complete history for Socrates must include his being a human being, assuming of course that 'human being' is a

sortal, and moreover his being a human being throughout that possible history. For by the omnitemporality thesis, Socrates was a human being throughout his actual existence. Now by the overlap requirement, any possible complete history for Socrates must include some segment of Socrates' **actual** history. But since **every** segment of Socrates' actual history includes Socrates' being human, **every** possible history for Socrates will include an actual segment in which he is human. And since 'being human' is a sortal, then by the omnitemporality thesis Socrates is human for his **entire** existence in **each** of the histories that are possible for him. (See Mackie P 1994, p.317.)

Mackie does not address the omnitemporality thesis. I shall do so further on. What Mackie does show is that the overlap requirement is too restrictive for essentialist purposes, and that even modified, it remains unacceptable.

An immediate problem for the overlap requirement is posed by properties such as *being born in Athens* and relational properties such as *living in a world that contains the pyramids*. For if Socrates was actually born in Athens, then he bears this property throughout his existence, so that by the overlap requirement he will have this property in every possible history. It then becomes essential to Socrates that he was born in Athens, which conclusion, as Mackie observes, is unwelcome to essentialists. By a similar line of reasoning, it will be essential to Socrates that he lives in a world containing the pyramids. (Mackie P 1994, pp.317-318.) But even if such properties can be exempted in some non *ad hoc* manner from the overlap requirement, there is the further problem that if Socrates **never** actually had a mole on his left foot, then by the overlap requirement there is no possible history for Socrates in which he did have a mole on his left foot **throughout** his existence. At best Socrates could have had such a mole only for most of his existence. Mackie

finds this too restrictive. (Mackie P 1994, p.318.)

Mackie therefore suggests a modification of the overlap requirement to the effect that if some history h is a possible complete history for an individual x, then either there must be some segment of h exactly the same as some segment of x's actual history h_a, or h must share its original segment with some other possible complete history that itself shares some segment with h_a. (Mackie P 1994, p.319.) This modification still secures the essentiality of sortals and allows moreover that even if Socrates never actually had a mole on his left foot, he could have had such a mole throughout his existence. But this comes at a price. Socrates could have had a mole on his left foot throughout his existence (history h) only because another history for Socrates is possible,

> one in which, although Socrates comes into existence with a mole on his left foot, the traceless removal of the mole is followed by his becoming exactly like the actual Socrates at some later stage. [history h#.] This is ludicrous. Perhaps there will always **be** such a possible history. But the idea that it is required to underwrite the possibility that a Socrates who never had a mole on his foot should have had one throughout his existence is absurd. The reason that we can envisage the possibility that Socrates should have had such a mole is, rather, that the precise distribution of moles about Socrates' person is un utterly trivial feature when it comes to envisaging alternative possible histories for Socrates. (Mackie P 1994, p.320.)

I am not sure that the notion of a property that is utterly trivial from a counterfactual perspective is all that clear, certainly not clear enough to justify Mackie's claim of ludicrousness. Nonetheless I think that the essentialist would take her point that the idea that what underwrites the possibility of h is the possibility of h# is unacceptable. The point presumably is that even if there are conditional possibilities, possibilities about moles shouldn't have to depend on other possibilities in the way that the modified overlap requirement demands. This completes our discussion of Brody's abortive attempt to establish sortal

essentialism.

4.8 WIGGINS ON THE ESSENTIALITY OF SORTALS

We turn now to consider Penelope Mackie's highly instructive discussion of Wiggins' notoriously obscure arguments for sortal essentialism. In chapters 2 and 3 of Wiggins' 1980 *Sameness and Substance*, Wiggins gives an account of sortals on which sortals are those general terms which provide principles of individuation (criteria of identity), principles which determine answers to questions about identity and distinctness at a time and over time. Thus for example there is supposedly associated with the sortal 'frog' a principle that determines the answers to questions such as whether a given portion of matter at a given time constitutes exactly one frog, or more or less than one frog; whether a frog at one time is the same frog as a frog at another time; whether spatio-temporal continuity is required for the persistence of a frog; whether the relation between a live frog and its corpse is that of identity; and in general, what sorts of alterations are compatible with the survival of an individual that is a frog. (See Mackie P 1994, p.324.)

I think that the textual evidence suggests that Wiggins would accept that the extension determined by my criterion of sortalhood (namely independent unity) coincides approximately with the extension determined by his criterion (namely principles of individuation). So I will proceed on the working assumption that the examples of sortals which I have hitherto referred to on my criterion also qualify as examples on his criterion. Now as Mackie observes (Mackie 1994, pp.321-324), according to Wiggins, an individual's principle of individuation is supplied by the **ultimate sortal** it falls under, and an individual cannot have incompatible principles of individuation at the same time. (See Wiggins 1980, chapters 2 and 3.) And of

course Wiggins will not allow that an individual change its principle of individuation for an incompatible principle of individuation, henceforth PI, over time. For a PI defines the temporal boundaries of an object's career, it determines what is to count as mere alteration and what as substantial change. And substantial change means extinction, so that if an object swapped one PI for another incompatible PI it would both have gone out of existence and yet be in existence, which is absurd. So objects cannot change their PI for an incompatible PI

While an object can fall under only one **ultimate** sortal, say, 'dog', it may simultaneously fall under more than one **non**-ultimate sortal, say, 'terrier' and 'fox terrier'. But then the cosatisfied sortals must have compatible PIs. That the PIs can be compatible without being identical is evinced by the consideration that one of these cosatisfied sortals may be more general; for example 'feline' is more general than 'lion', 'dog' is more general than 'terrier', and 'terrier' is more general than 'fox terrier'. And the synchronic and diachronic identity conditions for a feline would appear to differ from those for a lion, though they will be compatible.

Now according to Wiggins, an object has its PI essentially. This is provided by its ultimate sortal. (Wiggins 1980, p.122.) So it follows that an object satisfies its ultimate sortal essentially. Hence we have sortal essentialism. So Socrates was essentially human. Thus Socrates could not have been a centipede or a clay pipe or the squre root of pi, in so far as the PIs associated with these sortals are incompatible with the ultimate sortal 'human'. (Mackie 1994, p.322.)

Mackie suggests that it is consistent with the essentiality of PI that a lion could have passed its entire existence as a feline of a different kind, say, a leopard. (Mackie P 1994, p.323.) But this would follow in Wiggins' theory only if 'feline' rather than

'lion' is an ultimate sortal. This point brings us to the general issues of how we are to decide which of the sortals an object satisfies is its ultimate sortal, and what PIs are associated with what sortals, ultimate and non-ultimate. On this point Wiggins is representative of the literature in general: he refrains from explicit suggestions. So it is not at all clear whether cats share their PI with tigers or dogs or humans or frogs or butterflies or amoebas or even with fellow cats. Mackie points out how for his part Dummett thinks that all vertebrates have the same PI. (Mackie P 1994, p.329.) But unless we can decide these questions, we have no way of knowing whether a cat could not have been a butterfly throughout its existence. (See Mackie P 1994, pp.325-326.) A further complication in Wiggins' account is that he does not think that arifactual sortals are subject to the same kind of account as natural kind sortals.

Putting aside the problem of deciding the priority of sortals and the problem of determining the exact content of PIs, what reason is there to believe that an object has the PI associated with its ultimate sortal essentially? Mackie adduces and dismisses two arguments by Wiggins for the essentiality of PI.

The first is that counterfactual possibilities for actual individuals have to be anchored in the actual characteristics of those individuals. Otherwise we lose all ties with the object. (Wiggins 1980, pp.105-106.) This sort of consideration motivates Brody's overlap model of counterfactual possibilities. And Brody's model would secure the essentiality of PIs. But since Brody's model must be found wanting by the essentialist, what other reason is there to think that if the anchorage of an object must involve its actual characteristics then these would have to include its actual PI? As Mackie asks, why shouldn't we be able to anchor an object by the property of being a material object, even if Wiggins thinks that this is compatible with different

PIs? (Mackie P 1994, p.327.)

The second argument (Wiggins 1980, pp.105, 115-117) is as follows:

> A set of properties can provide the anchorage for an actual individual *x* in a counterfactual situation only if it includes properties sufficient to *individuate x* in the actual situation. But to individuate *x*, we must classify it under a sortal that supplies a ... [PI], a principle that determines its conditions for identity at a time and over time. So a thing's ... [PI] is essential to it....
>
> [But as Mackie notes, this] ... version of the anchorage argument begs the question. Why should we accept that a set of properties cannot anchor *x* in counterfactual situations unless the properties 'individuate' *x* in such a way as to determine its identity conditions at a time and over time? Because otherwise we cannot make sense of its being the *same* object in the counterfactual situation? Whether or not this is so is precisely the point at issue. (Mackie P 1994, p.328.)

Now it seems true that many people will pre-reflectively deny the possibility of a human being's having been a clay pipe or a crocodile. And this leads us to the argument that the idea that, say, a cat could have been a butterfly is ridiculous, for what could have made it the same individual if it had been as different as that? (Mackie P 1994, p.328.) Mackie concedes that perhaps the idea is ludicrous, but argues that the argument is insufficient to support the essentiality of PIs. For one would have to show that the reason that the supposition is ridiculous is that we are trying *per impossibile* to envisage the cat as having a different PI from its actual one. (Mackie P 1994, p.328.)

Nonetheless Mackie overlooks at this very point an argument for sortal essentialism that bypasses an appeal to the essentiality of PIs. For does not the supposed pre-reflective absurdity of trying to imagine a cat as a butterfly or a woman as a laurel or an oak tree as a cabbage point *prima facie* at least to sortal essentialism as an explanation? Cats are as different from butterflies as things can be. So what could

have made a cat the same thing as a butterfly? Or an oak tree a cabbage? Now Anscombe suggests that

> The doctrine that individuals have nothing that is essential to them suggests a phantasmic notion of the individual as a 'bare particular' with no properties, because it supposes a continued identity independent of what is true of the object. (Anscombe 1964, p.70.)

But it is not clear, *pace* Anscombe, that the denial of the existence of essential properties does commit one to the existence of *entia non grata* such as bare particulars or bare substrata. The idea is supposed to be that if an individual can survive over time or across worlds the loss of any property whatsoever, it must be something independent of any property whatsoever; hence it must be a bare particular. (See Vision 1970, p.322.) But as Vision points out, one can interpret the claim that it must be something independent of any property whatsoever in two ways: as the claim that it is not necessary for any individual to have any properties at all, or as the claim that there is no single property that is necessary for the identity of an individual. (Vision 1970, p.322.) All that the denial of essential properties commits one to is the latter interpretation, and so avoids the problem of bare particulars. Therefore we can dismiss as unfounded Forbes' suggestion that if a certain oak tree could have been a cabbage then we must have

> the conception of a subject which exists in two worlds, and is an oak tree in one and a cabbage in the other. ... and that such a conception is unattractively close to the scholastic notion of a substance as the **bare substratum** [my emphasis] in which an individual's properties inhere. (Forbes 1981, p.33.)

That is to say that we can deny the essentiality of a particular oak's being an oak without embracing bare particulars. This is because in general we can deny the essentiality to an individual of any of its properties without embracing an ontology of bare particulars.

Nevertheless, interpolated in the passage from Forbes just quoted there lies another consideration that suggests the essentiality of sortal properties. Forbes urges that we in fact lack a necessary condition of the possibility of the one object's satisfying the incompatible sortals 'oak' and 'cabbage', namely a conception of the one subject which exists as an oak tree in one world and a cabbage in the other. Forbes thinks th at we have no such conception. Wiggins seems to concur when he speaks of the conceptual cost involved in the admission of the possibility that one thing could have been anything we like. (See Wiggins 1980, p.214). If Forbes is not to beg the question, then he cannot be understanding 'conception of the one subject' in the sense of 'possibility of the one subject'. Thus he must mean something like conceivability or imaginability. And it is notoriously unclear whether what is conceivable is a reliable guide to what is possible, either logically or metaphysically. In any case I intend not to pursue this awkward question here, for if we don't in fact have a conception of the one subject existing as an oak in one world and a cabbage in another, and this is sufficient to render it impossible for an oak to have been a cabbage, then I think that equally we don't have a conception of the one object's having had all the properties of another and yet retaining its identity; so it is impossible for one object to have had all the properties of another. But since by Chisholm's argument essentialism commits one to the denial that this is impossible, essentialism is false, i.e.it is false that the distinction between essential and accidental properties is coherent. So Forbes' argument for sortal essentialism is self-refuting, in that it leads to the rejection of essentialism. (As for the question of bare particulars, far from anti-essentialists' being committed to an ontology of bare particulars, it seems that essentialists can **plausibly** deny the impossibility in Chisholm's argument only by an **appeal** to bare particulars. (See p.86 of this thesis.)

4.9 THE OMNITEMPORALITY THESIS

It is important to distinguish between on the one hand, entertaining the possibility of an **oak** tree's having come into existence as a **cabbage** and having then remained a cabbage for its **entire** existence, and on the other hand, entertaining the possibility of an oak tree's having **become** a cabbage at some point in its existence after it had already **come into** existence as an **oak** tree. The upshot of my discussion so far is that there is no good reason to think that an oak tree couldn't have spent its entire existence as a cabbage, at least no good reason that is not also a reason to reject essentialism. There is no good reason to believe that the one sortal property is essential over time to its bearer in **every** possible world in which that bearer exists. Now we recall from our discussion of Brody that an important feature of his argument for sortal essentialism was his appeal to the omnitemporality thesis, the idea that once an F, then at all times an F, where 'F' refers to a sortal. The omnitemporality thesis therefore excludes the second of the two possibilities enumerated at the beginning of this paragraph. (White suggests that Aristotle was concerned with the second possibility rather than with the first. (White 1972-73, p.69.)) I turn now to consider the omnitemporality thesis, the assumption widely held among essentialists that an object that is sortally F at one moment of its existence is so at every moment of its existence. I want to examine whether an object can change its sortal in the course of its career. If it in fact can't, then a sortal property will in effect be essential **over time** to its bearer **in the one possible world**, in as much as if the object had lost that property it would have gone out of existence in **that** possible world. But note that this will be compatible with a different sortal property's being essential over time to the same bearer in **another** possible world. This compatibility follows from my negative assessment of the view that the one sortal property has been shown to be essential over time to its bearer in

all possible worlds in which the bearer exists. So even if the omnitemporality thesis proves to be true, the essentialist would still have to allow that an object could have been, say, a bottle in one possible world for its entire existence and a butterfly in another possible world for its entire existence. This is a considerably more moderate form of sortal essentialism than is usually espoused. Now while I shall concede that there is a sense in which an object cannot change its sortal in the one possible world, I shall argue that the only reason that an object cannot do so arises purely from considerations of linguistic convention and so is of no **metaphysical** significance. The essential/accidental distinction fails to secure a foothold here.

Speaking now of sortals in terms of my criterion of independent unity (see p.207 of this thesis), we recall that according to the omnitemporality thesis, sortals are such that any instance x of a sortal S is such that if x instances S at any moment of its existence then it instances S at **every** moment of its existence. On Wiggins' criterion of sortalhood, namely providing a principle of individuation, it is also the case that the omnitemporality thesis holds good, with one proviso to be explained shortly. Now Wiggins can even **explain why** the omnitemporality thesis holds. For on his account, sortals provide the principles of individuation, and an object cannot change its principle of individuation for an incompatible principle of individuation. An object cannot change its PI, for a PI defines the temporal boundaries of an object's career, it determines what is to count as mere alteration and what as substantial change. Now substantial change means extinction, so that if an object swapped one PI for another incompatible PI it would both have gone out of existence and yet be in existence, which is absurd. So objects cannot change their PI for an incompatible PI. Therefore objects cannot change their sortal for a sortal with an **incompatible** principle of individuation. This is the proviso I referred to eleven lines above. And as it stands, it is consistent with an object's changing its

sortal provided that the new sortal has the same PI as the old sortal. And of course this raises the spectre of PIs all over again, with the attendant problem of deciding whether two sortals have the same PI. Unless we can resolve this issue we have no grounds for affirming on the Wigginsian criterion of sortalhood that no object can change its sortal.

Let us therefore pass by the Wigginsian criterion and its reference to the nebulous concept of a principle of individuation. Still there is something in Wiggins' account of sortals that the sortal essentialist could profitably salvage. That is the notion that a sortal is meant to define the **temporal** boundaries of the career of an object falling under that sortal. On this view, an object can survive sundry changes, but once it loses its sortal property, it ceases to exist. Can we justify this idea? Obviously we shall have to distinguish between phase sortals and non-phase or ultimate sortals. The idea is supposed to be that I can survive the lapse of my boyhood, so that 'boy' is a phase sortal subsumed under 'human being'. In contrast I am not supposed to survive the loss of my humanness, so that 'human being' is the ultimate sortal for me. Now I do not wish to deny that we pre-reflectively do tend to withhold identification of, say, an object that exits the morgue in a hearse with the powdery object that exits the crematorium in an urn. We might even generalise and say that no human can survive as a heap of ashes. But what grounds this widespread pre-reflective practice? Is there something of deep ontological significance here? To answer this question we need to recall my criterion of sortalhood: independent unity. We recall from the beginning of this chapter that we are interested in understanding authentic objects as opposed to inauthentic objects, and that the notion of an authentic object involves the notions of authentic unity at a time and over time. I said that both kinds of unity are analytically interdependent, and that the notion of causality helps to distinguish authentic unities from inauthentic unities,

diachronic or synchronic. Now my point is that while there is a fairly **non-arbitrary** distinction to be drawn between that portion of reality that constitutes an authentic object **at a time**, and that which constitutes the rest of the world at the same time, in contrast what counts as the **beginning** of the one authentic object's career and the **end** of the one authentic object's career are less non-arbitrarily and thus **more arbitrarily** decided The existence of phase sortals is a case in point. For why is a seedling thought to constitute a phase in the career of the one object (namely a grass) but the subsequent stalk of straw is not? Why for that matter isn't 'seedling' an ultimate sortal in its own right? Again, why doesn't the following series comprise the successive phases of the career of one authentic object: seed/seedling/grass/straw/compost/peat/coal/diamond? Answers cannot be found in considerations of spatio-temporal continuity, for the straw is as continuous with the grass as the grass is with the seedling. Nor will considerations of qualitative discontinuity help, for the butterfly is not less qualitatively dissimilar to the pupa than is a diamond to the coal. This leads me to suggest that our determination of the **temporal** boundaries of authentic objects is governed by stipulation and convention. In other words there is always some latitude for decision as to which behaviours to count as manifestations of an object that is sortally S and which to count as manifestations of an object that is not S but S_1.

Consider a sample of water which is evaporated, then condensed, then frozen. Commonsense practice is to say that the steam and the ice are just different **states** which the **one** quantity of water can assume, rather than that the steam and the ice are different **objects** corresponding to these different states. Now one might think that modern science can justify this practice, for science tells us that the same **molecules** survive the change of state, it is just that in evaporating (condensing/freezing) these molecules become more (less) spatially separated. Take

another example, this time an element such as a piece of copper. Brody suggests that the copper can change its colour when subjected to strong heat and yet remain the same piece of copper, but that if enough heat is supplied so that the copper is burnt, i.e.oxidised, then we no longer have the same object, viz. a piece of copper, but instead a compound composed of both copper and oxygen atoms. (Brody 1967, p.436.) The suggestion is implicit that **science** can explain why we have substantial change or mere alteration. In contrast, Teller suggests that when a beta particle separates from the nucleus of a carbon atom - an event which is then described as the transmutation of a carbon atom into an atom of nitrogen - this event

> could be glossed by saying that one existent, the carbon atom, ceases to exist and something else, a nitrogen atom, comes into existence; or by saying that a carbon atom changes into an atom of nitrogen while retaining its numerical identity. On the one hand, we could assimilate such transmutation to the case in which a table is chopped up into fire wood (and so ceases to exist); or we could assimilate it to the case in which a person simply loses an arm or leg ... It is not clear that contemporary atomic physics is committed to either of these descriptions, and it is also not clear that which of these descriptions is applied makes any difference to atomic physics. (Teller 1975, p.240.)

Teller draws the moral that how one glosses events such as transmutation depends on human needs, wants and interests. (Teller 1975, p.243.) He goes on to observe that if an object changes so that it no longer has one of those properties which are fundamental to the explanation of its other transient properties, i.e. if it no longer has its real essence, then

> so many of its properties and causal relations with other things change that one is inclined no longer to reidentify it as the same thing. In the example of an atom of carbon transmuted into nitrogen, the change in terms of mere number of protons and neutrons might be small. But the resulting change in terms of the properties of the resulting object is enormous. (Teller 1975, p.244.)

Teller seems to suggest that in the end it is the unity of the real essence underlying an object that determines the temporal boundaries in an object's career. But whether the reulting change is considered enormous enough to qualify as a case of detruction/creation as opposed to mere alteration is a matter for human decision.

Thus we judge the carbon to have gone out of existence and the nitrogen to have come into existence, but we do not judge one insect to have gone out of existence and another to have come into existence when the caterpillar becomes the pupa or the pupa the moth. In both cases the changes surely are enormous. The point is `that the unity over time of objects understood as **continuants** must be understood in Butler's **loose sense** of identity as opposed to Butler's **strict** sense of identity. The strict sense is too restrictive for ordinary purposes, for in the strict sense it is probably only atomic or subatomic entities that are ever identical over time. We reconcile identity over time with change because we appeal to a **loose** sense of identity. This does not mean however that the relation of loose identity does not conform to Leibniz's principle of the indiscernibility of identicals. For as Hirsch points out,

> if it is 'loosely' correct to say that a certain tree that exists at one time t is identical with a certain tree that exists at another time t1, then it must also be 'loosely' correct to ascribe to the former tree all the properties ascribed to the latter. Of course this does not mean that the tree has the same properties at t that it has at t1; rather, any properties possessed at a given time by the tree that exists at t are possessed at the given time by the tree that exists at t1. (Hirsch 1995, p.233.)

But then the **real essence** of the one object at time t is usually at best only **loosely** identical with the real essence of the same object at another time t1. For as an object's transient properties change, so will its real essence tend to have changed. The real esence need not necessarily change to account for a change in the object's derivative properties, for some cases of change in an object's derivative properties are due not so much to a change in its real essence as to changes in the interacting environment - recall the heating of the copper and its attendant temporary discoloration. But in so far as real essences determine the derivative properties, there will be many situations where changes in derivative properties will be due to changes in real essence. So real essences are transient too, but perhaps less so than

derivative properties as a matter of empirical fact. In any case it is this transience which calls for a loose sense of identity in the case of the real essence no less than in the case of the continuant as a whole - the real essence is a **part** of this whole. But then it is a matter for decision when changes in real essence are considered significant or enormous enough to count as the destruction of the real essence and thus of the continuant having that real essence.

The upshot of this discussion is that the temporal boundaries of the career of an authentic object are to some extent subject to negotiation. Thus in so far as these boundaries coincide with those associated with the conditions of application of a sortal, it is up to negotiation to some degree what to count as within the compass of a sortal and what not. Once the stipulation has been fixed, then it follows by logic that an object cannot lose its ultimate sortal, for the object's career is defined in terms of this sortal. To this extent the sortal **is** essential to the object's identity. But this is merely conceptual essentialism. It is a far cry from the claim that an object could not in some **metaphysical** sense have lost its ultimate sortal. The reason why an object could not have lost its ultimate sortal has nothing to do with metaphysics; rather the reason is one of linguistic convention: the object ceases to exist upon losing its sortal because that is the way we have decided to trace the object's career. The omnitemporality thesis is a thesis about how we have decided to use language, and not a thesis about how the world **must** work.

In this connection Copi observes that the alleged distinction between essence and accident

> is simply a projection of differences in human interests or a reflection of peculiarities of vocabulary. ...
> The distinction between different kinds of change ... is subjective rather than objective. We happen to be interested, usually, in some attributes ... more than in others. When ... [change occurs], we say that ... [the same thing] persists through the change provided that it does not lose those attributes by whose possession it satisfies our interests. For example, our interest in tables

> is for the most part independent of their colors. Hence that interest remains satisfiable by a given table regardless of any alteration it may suffer with respect to color. Paint a brown table green, and it remains substantially or essentially the same; the change was only an accidental one. If our interests were different, the same objective fact [of a brown table being succeeded by a green table] would be classified quite differently. Were our interest to lie in **brown** tables exclusively, then the application of green paint would destroy the object of our interest, would change it substantially or essentially from something which satisfied our interest to something which did not. The implication is that attributes are neither essential nor accidental in themselves, but can be so classified only on the basis of our subjective interests in them. ... That we regard a table as essentially the same despite alteration in color or movement from place to place is a consequence of the peculiar nature and limitations of our vocabulary, which has a single word for tables, regardless of color, but lacks special words for tables of different colors. Suppose that our language contained no word for tables in general, but had instead - say - the word 'towble' for brown table and the word 'teeble' for green table. Then the application of green paint to a towble would be said to change it essentially, it might be argued, for no towble would remain; in its place would appear a teeble. Or if there were a single word which applied indiscriminately to tables and heaps of ashes, say, 'tashble', with no special substantive denoting either of them univocally, then perhaps the destructuion of a table by fire would not be regarded as an essential change. That which appeared at the end of the process would admittedly be in a different state from what was there at the start, but it would still be identifiably the same tashble. [Compare the case of ice/water/steam.] (Copi 1954, pp.335-336.)

I would like to close this discussion of the omnitemporality thesis with some remarks based on some delightfully insightful observations by Margolis in his 'Dracula the Man: An Essay in the Logic of Individuation'. (see Margolis 1964, p.541ff.) Now as a matter of empirical fact the career of a human being has never included phases where the humanoid shape disappears intermittently to be replaced by a bat-like shape. The world is just not governed by that sort of causal regularity, at least not outside fiction. But if there did prove to be (frequent) instances of a spatio-temporally, qualitatively and causally continuous series comprising humanoid shapes succeeded by bat-like shapes, then how we would delimit the individuative boundaries of such a series would be a matter for decision. Firstly, we could refuse to identify the count with the bat on the grounds that the count is a man (male human) and men can't be bats. In that case we would be committed to the intermittent existence of the Dracula the man; but that would be all right, since after

all we allow the intermittent existence of watches through disassembly and reassembly. Secondly, we could revise our criteria for being a man and rule that 'man/male human' is no longer an ultimate sortal but a phase sortal like 'boy' or 'teenager' or 'adult'. Indeed the precedent is already set by sex changes. Similarly, we could rule that 'bat ' is no longer an ultimate sortal but a phase sortal. On this view, we coin a new ultimate sortal, viz. 'vampire', and rule that being a man and being a bat are recurrent phases in the career of a single individual, namely a vampire, which individual has wings *qua* bat but lacks them *qua* man. Or thirdly, we could adopt a position that combines elements of the previous two. We could invent a new ultimate sortal, 'vampire' and recognise the phase sortals 'vampman' and 'vampbat'. We understand 'man' and 'bat' in their present ordinary senses. On this third view, a vampire *qua* vampman is indistinguishable from a man, and a vampman *qua* vampbat is indistinguishable from a bat. Nevertheless we cannot predicate 'man' of a vampman, for that would be akin to saying that a boy is a dog, which is impermissible on the ordinary criteria of 'boy' and 'dog'. Nor can we predicate 'bat' of a vampbat, for that would commit the same offence to usage. The point is that how we decide **is** just a matter for decision. This is not meant to cast doubt on the objective existence of a state of affairs in which humanoid phenomena are succeeded by bat-like phenomena. Nor is it to question the intuition that these disparate phenomena consitute authentic units at a time. All it is meant to do is to highlight that the exact location of individuative boundaries in the series is less than fully objective. So while it follows by definition that an object cannot lose its ultimate sortal, the sortal which sets its temporal limits, there is nothing profound philosophically in this. The omnitemporality thesis is a thesis about the essential properties of objects which **we** choose to stipulate, and not which **reality** gives to us.

The upshot of my discussion in this chapter is that there is no cogent case to be made for the essentiality of sortal properties.

GLOSSARY

(METAPHYSICAL) ESSENTIALISM: the view which endorses the coherency of the distinction between those properties which are metaphysically essential to an object and those that are not essential (i.e. they are accidental). (See p.7.)

ESSENTIALIST (noun): a proponent of essentialism. (See p.7.)

ESSENTIALIST (adj.): pertaining to essentialism.

ANTI-ESSENTIALIST (noun): an opponent of essentialism; (adj.) pertaining to opposition to essentialism. (See p.8.)

FULL-ESSENTIALISM: the view that some objects have essential properties. (See p.7.)

FULL-ESSENTIALIST:a proponent of full-essentialism or pertaining thereto. (See p.8.)

ANTI-FULL-ESSENTIALIST: an opponent of full-essentialism or pertaining thereto. (See p.8.)

CATHOLIC ESSENTIALISM: the view that all objects must have the same essential properties. (See p.7.)

INDIVIDUALIST ESSENTIALISM: the view that not every one of a particular object's essential properties need be essential to another object. (See p.7.)

WEAK ESSENTIALISM: the thesis that some properties are essential to every object. (See p.19.)

STRONG ESSENTIALISM: the thesis that either some properties, while not possessed by every object, are possessed essentially by all their bearers, or that some properties are possessed essentially only by some but not all of their bearers. (See p.19.)

LEIBNIZIAN ESSENTIALISM: the thesis that every one of any object's properties is essential to that object. (See p.19.)

CONCEPTUAL ESSENTIALISM: the view that a necessary *de re* statement is

necessary in virtue of the logical entailment of the predicate by the subject term designating the **concept** of the statement's subject. (See p.17.)

THE PARITY THESIS: a name is used referentially or has referential force just where its semantic function is to refer; its sense, if it has one, is semantically idle. The parity thesis is the thesis that where 'a' is a non-empty name with **referential force**, it is necessarily true **that** a is F if and only if it is necessarily true **of** a that it is F. (See p.24.)

THE SEMANTIC PRECEDENT: '(necF)a' is to be true on an interpretation I just when the referent I(a) of the individual constant 'a' is such that it is necessarily true of it that it is F. (See p.27.)

THE POSSIBILITY PRINCIPLE: an object x could have had exactly the genuine properties of an object which differs from x only in the circumstance that one of x's genuine properties is replaced by another not possessed by x. (See p.79.)

THE PRINCIPLE OF THE TRANSITIVITY OF PREDICATIVE POSSIBILITY: if an object a that is R could have had exactly the properties of an object that is predicatively S, then if an object that is predicatively S could in turn have had exactly the properties of an object that is predicatively T, then a could also have had exactly the properties of an object that is predicatively T. (See p.80.)

PIOICP THE PRINCIPLE OF THE IDENTITY OF THE QUALITATIVELY INDISCERNIBLE IN CASES OF COUNTERFACTUAL POSSIBILITY: if **actual** a could have had all of **actual** b's properties, then a could have been b. (See p.81.)

LNI THE NECESSITY OF DISTINCTNESS: the thesis that if a and b are distinct, then they are necessarily so. (See p.89.)

LI THE NECESSITY OF IDENTITY: the thesis that if a and b are identical, then they are necessarily so. (See p.90.)

HAECCEITY: the property of being identical with Socrates is Socrates' haecceity.

(See p.93.)

HAECCEITISM: the thesis that an object's haecceity is a genuine and non-composite property of the object. (See p.103.)

PREMISE I: It is impossible for the same table x to be originally constructed from a hunk of matter y and in addition to be simultaneously originally constructed from a distinct hunk of matter y*. (See p.163.)

PREMISE IV*: For any possible table x and any possible hunks of matter y and y*, and any plan P, if it is possible for x to be originally constructed from y according to plan P, where y* **does not overlap** with y, and it is also possible for a table to be originally constructed from y* according to plan P, then it is also possible [both] that x be originally constructed from y according to P, ... [while] in addition ...that [simultaneously] some table or other be originally constructed from y* according to P. (See p.162.)

PREMISE V*: If it is possible for a table x to be originally constructed from a hunk of matter y according to a certain plan P, then necessarily, any table originally constructed from hunk y according to the same plan is the very table x and no other. (See p.163.)

PREMISE (V)**: if it is possible for a table x to be such that it is originally constructed from a certain hunk of matter y according to a certain plan P, **and** such that x is the **only** table so constructed, then necessarily, any table that is the only table to be originally constructed from y according to P is identical with x. (See p.164.)

SALMON'S MODIFIED ARGUMENT: this argument is constituted by Premise V** together with premise (I) and a correspondingly strengthened version of

premise (IV*). (See p.164.)

NOONAN'S PROPOSAL: if it is possible for x to be the sole F to be originally constructed from a hunk of matter y according to plan P at a certain time t, then necessarily any F originally constructed from y at t according to P is x. (See p.167.)

AMENDED VERSION OF NOONAN'S PROPOSAL: if in a world W accessible from a world w, x is the only F to be originally constructed from a hunk of matter y according to plan P at time t, then in any world accessible from w any F originally constructed from y at t according to P is x. (See p.170.)

WC: WC is the claim that sameness of **all but a sufficiently small portion** of original matter plus sameness of thing kind plus sameness of moment of origin is sufficient for transworld identity. (See p.167.)

THE PRINCIPLE OF NO BARE IDENTITIES: the thesis that the difference between two qualitatively identical possible states of affairs cannot consist **solely** in the fact that they concern two distinct individuals a and b; i.e. if there are no differences in the identities of individuals other than a and b which ground the difference in question, then the two possible states of affairs cannot be qualitatively identical after all, if, that is, they are indeed to concern **distinct** a and b. (See p.176.)

THE PRINCIPLE THAT IDENTITY NOT BE DETERMINED BY EXTRINSIC FEATURES: the principle that intrinsic rather than extrinsic features should ground the facts of transworld identity and distinctness. An intended corollary is that it cannot be the case that a is identical with b just because nothing else is present to compete with a for identity with b. (See p.177.)

THE OMNITEMPORALITY THESIS: if an object is sortally S at one moment of its career then it is is sortally S at all moments of that same career. (See p.219.)

PI PRINCIPLE OF INDIVIDUATION: principles which determine answers to questions about identity and distinctness at a time and over time. (See p.222.)

BIBLIOGRAPHY

Abbreviations

A Analysis

AJP Australasian Journal of Philosophy

APQ American Philosophical Quarterly

ASP Aristotelian Society Proceedings

CJP Canadian Journal of Philosophy

CM *A Companion to Metaphysics* ed. Kim, J. and Sosa, E. (1995. Blackwell)

DYP Danish Yearbook of Philosophy

IPQ International Philosophical Quarterly

JP Journal of Philosophy

JPL Journal of Philosophical Logic

MXI *Midwest Studies in Philosophy v.XI. Studies in Essentialism*, ed French, P. *et al* (1986. Uni of Minnesota Press)

M Mind

N Nous

Ph Philosophia

PQ Philosophical Quarterly

PR Philosophical Review

P Philosophy

PPR Philosophy and Phenomenological Research

RM Review of Metaphysics

S Synthese

WP *The Ways of Paradox* Quine, W. (1976. Harvard Uni Press)

Adams, R. 1979 'Primitive Thisness and Primitive Identity', JP v.76 pp.5-26.

Anscombe, E. 1964 'Substance', ASP supp. v.38 pp.69-78.

Anscombe, E. 1981 *Metaphysics and the Philosophy of Mind* (Blackwell)

Aune, B. 1995 'Nature' in CM pp.349-350.

Ayer, A. 1975 'Identity and Reference', Ph v.5 pp.171-192.

Ayer, A. 1982 *Philosophy in the Twentieth Century* (Unwin Paperbacks)

Ayers, M. 1981 'Locke versus Aristotle on Natural Kinds', JP v.78 pp.247-272.

Bogen, J. 1966 'Identity and Origin', A v..26 pp.160-165.

Brody, B. 1967 'Natural Kinds and Real Essences' JP v.64 pp.431-446.

Brody, B. 1973 'Why Settle for Anything Less than Good Oldfashioned Aristotelian Essentialism?' N v.7 pp.351-365.

Brody, B. 1980 *Identity and Essence* (Princeton Uni Press)

Carter, W. 1983 'Salmon on Artifact Origins and Lost Possibilities', PR v.92 pp.223-231.

Carter, W. 1986 'Mapping Semantic Paths: Is Essentialism Relevant?' in MXI pp.53-73.

Cartwright, R. 1968 'Some Remarks on Essentialism' in Cartwright, R. *Philosophical Essays* (1987. The MIT Press)

Chisholm, R. 1967 'Identity Through Possible Worlds: Some Questions' in Loux 1979 pp.80-87.

Cohen, S.M. 1977-78 'Essentialism in Aristotle', RM v.31 pp.387-405.

Cook, M. 1987 'Difference at Origin', Ph v.17 pp.501-507.

Copi, I. 1954 'Essence and Accident' in Loux, M. (ed.) *Universals and Particulars* (1976. Uni of Notre Dame Press) pp.331-346.

Della Rocca, M. 1996 'Essentialists and Essentialism', JP v.93 pp.186-202.

Elliot, R and Gallois, A. 1984 'Would It Have Been Me? (Against the Necessity of Origin) AJP v.62 pp.292-293.

Forbes, G. 1981 'An Anti-Essentialist Note on Substances', A v.41 pp.32-37.

Forbes, G. 1983 'Thisness and Vagueness', S v.54 pp.235-259.

Forbes, G. 1985 *The Metaphysics of Modality* (OUP)

Gibbard, A. 1975 'Contingent Identity', JPL v.4 pp.187-221.

Gupta, A. 1980 *The Logic of Common Nouns* (Yale Uni Press)

Haack, S. 1978 *The Philosophy Of Logics* (CUP)

Hirsch, E. 1982 *The Concept of Identity* (OUP)

Hirsch, E. 1995 'Identity' in CM pp.229-234.

Hughes, G. and Cresswell, M. 1968 *An Introduction to Modal Logic* (Methuen)

Kripke, S. 1980 *Naming and Necessity* (Blackwell)

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

Loux, M. (ed.) 1979 *The Possible and the Actual* (Cornell Uni Press)

Lowe, E. 1989 *Kinds of Being* (Blackwell)

McGinn, C. 1976 'On the Necessity of Origin', JP v.73 pp.127-135.

Mackie, J. 1974 '*De* What *Re* is *De Re* Modality?', JP v.71 pp.551-561.

Mackie, P. 1987 'Essence, Origin and Bare Identity', M v.96 pp.173-201.

Mackie, P. 1994 'Sortal Concepts and Essential Properties' PQ v.44 pp.311-333.

Margolis, J. 1964 'Dracula the Man: An Essay in the Logic of Individuation', IPQ v.4 pp.541-549.

Mason, H. 1967 (ed. and trans.) *The Leibniz-Arnauld Correspondence* (Manchester Uni Press)

Matthews, G. 1990 'Aristotelian Essentialism', PPR v.50 supp pp.251-262.

Moravscik, J. 1976 'The Discernibility of Identicals', JP v.73 pp.587-598.

Mortensen, C. 1983 'Relevance and Verisimilitude', S v.55 pp.353-364.

Neale, S. 1990 *Descriptions* (The MIT Press)

Noonan, H. 1983 'The Necessity of Origin', M v.92 pp.1-20.

Odegard, D. 1975 'Locke's Unnatural Kinds', A v.35 p.208.

Parsons, T. 1969 'Essentialism and Quantified Modal Logic' in Linsky, L. (ed.) *Reference and Modality*, pp.73-87. (1977. OUP)

Plantinga, A. 1974 *The Nature of Necessity* (OUP)

Plantinga, A. 1980 *Does God Have a Nature?* (Marquette Uni Press)

Price, M. 1982 'On the Non-Necessity of Origin', CJP v.12 pp.33-45.

Prior, A. 1957 'Opposite Number' in *Papers in Logic and Ethics* ed. Geach, P. and Kenny, A. (1976. Duckworth)

Prior, A. 1959-1960 'Identifiable Individuals', RM v.13 pp.684-696.

Putnam, H. 1970 'Is Semantics Possible?' in Schwartz 1977 pp.102-118.

Putnam, H. 1975 'The Meaning of 'Meaning'' in Putnam, H. *Mind, Language and Reality* (1975. CUP)

Quine, W. 1953 'Reference and Modality' in Sleigh, R. (ed.) *Necessary Truth* (1972. Prentice-Hall) pp.135-151.

Quine, W. 1953 'Three Grades of Modal Involvement' in WP.

Quine, W. 1960 *Word and Object* (The MIT Press)

Quine, W. 1962 'Reply to Professor Marcus' in WP.

Quine, W. 1976 'Worlds Away' in Quine, W. *Theories and Things* (1981. Harvard Uni Press)

Quine, W. 1990 'Comment on Marcus' in Barrett, R. and Gibson, R. (ed.) *Perspectives on Quine* (1990. Blackwell) p.244.

Ravnkilde, J. 1975 'How to Get Rid of Essentialism', DYP v.12 pp.7-64.

Rescher, N. 1974 *Studies in Modality* (Blackwell. APQ Monograph Series.))

Sainsbury, M. 1991 *Logical Forms* (Blackwell)

Sainsbury, M. 1995 'Philosophical Logic' in Grayling, A. (ed.) *Philosophy A Guide through the Subject*, (OUP) pp.61-122.

Salmon, N. 1982 *Reference and Essence* (Blackwell)

Salmon, N. 1984 'Impossible Worlds', A v.44 pp.114-117.

Salmon, N. 1986 'Modal Paradox: Parts and Counterparts, Points and Counterpoints' in MXI pp.75-120.

Salmon, N. 1987 'The Fact that $X = Y$', Ph v.17 pp.517-518.

Sarker, H. 1982 'Origins and Identities', AJP v.60 pp.140-151

Schlesinger, G. 1994 *Timely Topics* (St. Martin's Press)

Schwartz, S. (ed.) 1977 *Naming, Necessity and Natural Kinds* (Cornell Uni Press)

Shoemaker, S. 1979 'Identity, Properties and Causality' in Shoemaker, S. *Identity, Cause and Mind* (1984. CUP)

Sober, E. 1995 1995 'Natural Kinds' in CM pp.345-346.

Sprigge, T. 1962 'Internal and External Properties', M v.71 pp.197-212.

Stalnaker, R. 1979 'Anti-Essentialism' in *Midwest Studies in Philosophy v IV Studies in Metaphysics* ed. French, P. *et al* (1979. Uni of Minnesota Press) pp.343-355.

Teller, P. 1975 'Essential Properties: Some Problems and Conjectures', JP v.72 pp.233-248.

Van Inwagen, P. 1990 'Four Dimensional Objects', N v.24 pp.245-255.

Vision, G. 1970 'Essentialism and the Senses of Proper Names', APQ v.7 pp.321-330.

White, N. 1972-73 'Origins of Aristotle's Essentialism', RM v.26 pp.57-85.

Wiggins, D. 1974 'Essentialism, Continuity and Identity', S v.23 pp.321-359.

Wiggins, D. 1980 *Sameness and Substance* (Blackwell)

Wilkerson, T. 1988 'Natural Kinds', P v.63 pp.29-42.