US
University of Sussex

A Study in Metaphysics for Free Will

Using models of Causality, Determinism and Supervenience

in the search for Free Will

David Robson

Submitted for the degree of Doctor of Philosophy

University of Sussex

September 2013

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

David Robson

UNIVERSITY OF SUSSEX

David Robson

Doctor of Philosophy

A Study in Metaphysics for Free Will

Summary

We have two main aims: to construct mathematical models for analysing determinism, causality and supervenience; and then to use these to demonstrate the possibility of constructing an ontic construal of the operation of free will - one requiring both the presentation of genuine alternatives to an agent and their selecting between them in a manner that permits the attribution of responsibility.

Determinism is modelled using trans-temporal ontic links between discrete juxtaposed universe states and shown to be distinct from predictability. Causality is defined on a temporal sequence of σ-algebras and quantified using a measure. The measure leads to definitions of causal over-determination and epiphenomena. Proofs are constructed to demonstrate deterministic universes must carry their properties essentially but not necessarily locally. We argue determinism and causality are separate doctrines.

These models and results are marshalled to put the case that a counterfactual construal of ontic choice cannot work. In response we propose 'immanence' - a modified form of indeterminism whereby a universe can present choices to its denizens.

We prove that beings subsumed within a universe cannot pilot their own actions. We then argue these beings can exercise free will only when selecting between choices inhering within immanent relata. A being is responsible for its selections if and only if it is constituted of a temporally evolving deterministic substructure. Our proposal is novel: it avoids injecting indeterminism into the decision process.

Topological models for property supervenience are developed and used to reconstruct standard definitions from the literature. These are then used to demonstrate considerations of supervenience do not affect our arguments.

We have demonstrated that a model of the exercise of free will involving both genuine choices and responsibility is possible but can only operate within a non-deterministic universe possessing specific traits.

Acknowledgements

I should like to thank Professor Michael Morris, without whom this thesis would not have been written. His mentorship, guidance and philosophical insights have been invaluable throughout. I am grateful for the philosophy he has taught me, the complications he has helped to unravel and the many hours he has spent discussing these ideas as they took shape. From him I have learnt the supreme importance of clear, methodical thinking and the need to express my ideas in the most transparent prose possible.

I am grateful to the University of Sussex for affording me the opportunity to undertake this study.

I am also grateful to Professor Basil Hiley. He has spent countless hours with me discussing and explaining the philosophical issues that plague the foundations of quantum theory. The generosity he has shown with his time and his willingness to explore new ideas to their foundations without allowing prejudice or preconceptions to skew the debate have shown me how the very best research should be done.

Over the years I have many conversations with colleagues and friends on philosophy and on issues in the foundations of physics. I have learnt so much from each and every one of them and I thank them all; in particular I should like to mention Bob Callaghan, Dr Owen Maroney, Dr Keith Bowden and Professor Rainer Zimmermann.

I dedicate this study to my parents, to Shaila Duwal and to the memory of Tulsi Man Duwal and Birendra Hada - each, in their own way, an extraordinary and inspirational person.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction: aims and scope

This study is ostensibly concerned with free will; but it is just as much concerned with building mathematically-couched metaphysical models. Free will acts as a sun around which we orbit as we construct models that allow talk of determinism, causality and supervenience - but talk with the idea of free will often glinting suggestively in the background. Throughout our aim is to attain an 'ontic' perspective on the issues at hand. An ontic view in as much as we aspire to produce an understanding of our metaphysics from a 'God's eye view' - so that we are largely liberated from subjective baggage. And ontic also because we adopt a realist stance: there is a world external to us in which we - our bodies, our thoughts - are subsumed as some small part.

Our aims are broadly three: first, we aim to present an [ontic] model of the exercise of free will that satisfies some basic requirements that we set out below. Second, we use our first aim to motivate the building of a set of mathematically styled models of causality, determinism and supervenience. We tackle these three notions as they have an important place in many debates that orbit around free will and they play a crucial part in our arguments. And third, we aim to lay a foundation for future work. This study is primarily concerned with the first two aims and the third, whilst lurking in the background, makes only a shadowy appearance here and there. We shall meet some hints of the direction the third aim aspires to travel scattered in the text and a little more explicitly in the concluding chapter.

Before we move to discuss the structure of this work in detail, we shall say a little more about free will.

What then is free will? If no one asks me, I know what it is. If I wish to explain it to him who asks, I do not know.[1] Free will is a tricky and slippery philosophical beast: it is not easy to pin

[1]With apologies to St.Augustine, XI, 14.

down and, once seemingly secured, it is often doubtful the pins are ever really robust enough to withstand the pickings of dissenting voices. And nor is it clear if we, as conscious beings, possess free will in any of its proposed guises. However, there is one thing we may say about free will and we hold to it here: to have free will is to be able to choose [and, perhaps, then to act] and to do so in a manner that makes us responsible for our selections [and actions]. What it is to be presented with choices and to choose in a manner that warrants the attribution of responsibility are what we are to investigate. We outline the precise criteria that we seek for free will below.

In the approach we adopt here we shall not worry too much about the move from selection to action. In simple terms, what we propose forces some action following choices but those actions play no crucial part in satisfying what we seek for the exercise of free will. We do not take this as an issue as our models look at structures and not the specific workings of the mind or body in any explicit manner. We take it that selections lead to appropriate actions that are, in the deeper realms that involve the minutiae of actions and thoughts, consistent with the choices that are made and with the mental movements that follow choices. Our model is one that looks on humans - as free agents or beings who can be responsible - in an external, sterile manner. Our models seek to look for mechanisms that might then act as a launch for a more detailed unpacking of issues that surround moves from responsibility to moral responsibility but we do not worry about morality here. The limited scope of our aims and models is a strength: it allows us to dig down to the basic machinations of humans as denizens of, and as sterile objects in, a universe; and it is also a weakness: it leaves an unsatisfactory feeling that something is missing. Our sterile beings are devoid of the richness of internal lives - internal lives where we might find many more layers to the idea of free will than our models let us reveal.

So, although often ‘responsibility’ is taken to mean ‘moral responsibility’, here we talk of responsibility and not moral responsibility. We do not dismiss the potential importance of the move from responsibility to moral responsibility in the debate but we seek only necessary ontic conditions, or perhaps a subset of necessary conditions, for the exercise of free will and this is not to include morality.[2] Rather we take moral responsibility both to be a nebulous aspect of - or perhaps even absent from - our models: they do not accommodate the machinations of the mental in any fashion that might be useful for morality. So morality is something our models do not, and probably cannot, accommodate. If our version of free will is lacking, the issue of moral responsibility is certainly one place where that lack is obvious. Whilst we shall make a few tentative comments on this here and there, it is not an issue that is to bother us unduly.

[2] For further discussion see, for instance, Frankfurt (1971), Frankfurt (1982), Frankfurt (1988), Frankfurt (1992) and Pettit (2001).

What we shall eventually seek for free will, as necessary conditions for its operation, are the ability to choose from amongst genuine choices and to be deemed responsible for our selections. These two pillars, choices and responsibility, as we loosely have it here, seem to clash when looked at more carefully: they curdle irreversibly when thrown into the same mixing bowl. It is this difficulty, we claim, that our model circumvents.

A way of unpacking free will which will help us understand this difficulty - the clash between choices, as alternative possibilities, and responsibility - can be set out as follows:
If I possess free will, then when presented with choices I am free to deliberate upon them and then choose one rather than the other.[3] I am then, presumably, able to translate my choice into action without undue constraint. And I can be taken to choose freely in the sense that were I to use some handy time machine, return to some moment in that past and start the process again, then I could, on deliberating again, choose what I previously rejected and reject what I previously chose. And my selection is the result of my piloting of my deliberations. And I am unconstrained in that my actions following my choices are consistent with the way the world is - the laws of nature. So I can jump up to reach an apple should I choose but I cannot jump up to reach the moon. Choices, alternative futures, must be limited to those that can be acted upon in most cases; or at least those permitted by the laws of nature. In simple terms, we can take it that this form of choice that we need for free will has at its base the truth of the maxim that 'I could have deliberated and then chosen to do otherwise.' We might say that that I am responsible for my unconstrained action if, although I chose A, I could have chosen B instead.[4] I kicked the ball and I am responsible for doing so.[5] I am responsible because I could have chosen not to kick the ball but rather I choose, after careful deliberation and thought, that I would kick it. As we shall explain in our later arguments, there is a difficulty at base with this rather crude encapsulation of free will: it cannot launch as it stands. Eventually we shall show deliberation can play no part and so the maxim 'I could have chosen to do otherwise' stands in opposition to any attribution of responsibility. Deliberation cannot be what we take it to be, it cannot be a concatenation of thoughts, or more widely mental acts, that are piloted by a self in a 'free' manner.

So what are the difficulties with this way of thinking about free will? Where is the clash of which we spoke? We start by noting that we are part of the universe, our thoughts, our bodies and our

[3]Later, in sections 5.6 and 5.7, we show that a process of deliberation followed by selection is impossible even under generous assumptions.

[4]Where A and B are not related, so that doing B does not entail A and vice versa.

[5]There are no great worries for moral responsibility here as kicking a ball is not a moral act. We could substitute 'kicking a cat', perhaps, if moral responsibility were to be more to the fore.

being all inhere within it.[6] Choosing to do otherwise than one did requires that the universe, of which we are a part, be able to follow different paths from some moment in its evolution. And it requires that we be able to pilot and select between these different paths through our deliberations, the very deliberations that lead to selections from choices. If we can follow more than one path from some moment then it would seem that the possibility of choosing otherwise must be eked out of this fact by careful analysis. But in eking out the possibility of alternative paths, we will find ourselves admitting that there must be some moment at which the universe turns one way rather than the other by pure serendipity. And when it does turn, we cannot attribute responsibility to any part of that universe, even to ourselves as beings deliberating within it, for the turn: there can be no move from serendipity to responsibility as we lurch uncontrollably, even momentarily, to the left or to the right. So it would seem that indeterminism, the doctrine that the universe may lurch hither and thither as it evolves without rhyme or reason, puts halt to attributing responsibility but is seemingly necessary for the possibility of alternatives.

Perhaps we could abandon indeterminism and seek refuge in determinism for resolution? If we take it that the universe is deterministic - broadly the doctrine that the future is fixed by the present and made inevitable by it - then we can identify the route through which our deliberations pass and then find some reason to say that this person or that person is responsible as selections and deliberations passed through them. But still we must face the difficulty that determinism throws up: it does not give us choice. What we do, our 'choosing', is fixed before we choose it. The choices we make are not 'up to us' any more, we cannot deliberate to choice in any meaningful manner. And if this is the case then we must ask what sort of responsibility are we to achieve if actions and the like merely pass through us on their way from the distant past to the selections we make and then onwards into the future? We cannot jump into our time machine, return to some past point, and deliberate to a different choice; rather the universe fixes it all for us. It is nicely put by van Inwagen:

> If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things [including our present acts] are not up to us. (van Inwagen, 1986, p. v)

This is the dilemma that any purported solution to the exercise of free will presents to us: it requires both the presentation of genuine choices - alternative futures - and the ability to deliberate and to select between them; and the machinations of the selection must justify the attribution of

[6] We unpack and justify this statement in more detail in section 5.5.

responsibility. But indeterminism seemingly destroys any possible move from selection to the attribution of responsibility.

The debates around these issues are extensive, involved and ingenious. There might well be as many solutions, rejections of parts of the arguments, reworking of ideas of freedom and the like as there are philosophers and possibly many more.[7]

If we were to summarise very briefly the main broad sweeps of the free will debate, we would start with dividing the protagonists into two broad camps: compatibilist and incompatibilist. Then note that these broad camps contain myriad subgroups and gradations of subtlety as to what exactly is meant by free will and how it fits into the proposed models.[8] Broadly compatibilists consider that we may have free will, in some form, even if determinism obtains. Within the bound of this are many possible ways of cashing out the details - for instance, Strawson (1984, p 5) says that compatibilists believe 'possibly both D and F' but then he explains that this has a wide scope of possibilities:[9] everything from 'D is not true, but that we are free, and would be free even if D were true' to 'D is true, and we are free.'[10] Incompatibilists deny that free will is possible in deterministic universes. Some incompatibilists, the Libertarians, go on to argue that it is possible in indeterministic settings. Others argue that such Libertarian contentions are incoherent or impossible.[11] There are also hard-determinists who take it that the universe is deterministic and that this means that we cannot be free.[12] And then there are soft-determinists for whom determinism actually obtains as does free will.

And there are those that seek to combine the best of both determinism and indeterminism into 'Valerian' solutions. These solutions attempt to gain leverage on choice by injecting indeterminism into an individual's deliberations - the deliberations that move from presentation of choices to choosing. There are a varied set of Valerian models proposed in the literature. Sometimes the indeterminism is injected at the beginning of the deliberation and in some models at the end of the deliberation. We reject models that leave indeterminism within the self as we argue they fail to provide an adequate lever for responsibility: we discuss this in Chapter 5.[13]

---

[7] Some recent surveys include Watson (2003), Fischer et al. (2007), Kane (2011) and Russell & Deery (2013).

[8] The best general summary of the technical positions adopted can be found in Strawson (1984, Chapter 1). The very brief summary of these gradations that follows owes much to Strawson's clear delineation and uses his terminology.

[9] D = determinism and F = freedom.

[10] Quotes from Strawson (1984, p 5).

[11] For instance, Strawson (1984), Double (1990), Honderich (2002), Smilansky (2002), Wegner (2003), Pereboom (2006).

[12] Both the terms 'hard-determinism' and 'soft-determinism' originate in James (2010).

[13] Examples of such Valerian theories can be found in Kane (1985), Double (1990), Dennett (1997), Mele (2001), and Clarke (2005).

And there are those that claim that 'being able to do otherwise' is not the route to take to a solution;[14] and those that claim that the debate itself has misconstrued what we should take as the essential issues.[15]

And, as a final note, there are claims that the debate need not concern determinism and indeterminism at all - such as Strawson (1984) - as they can have no relevance. Here, of course, we reject this rejection. Our rejection will be judged as well-judged only in the light of the success of our models in satisfying the criteria we set out for free will below. If our success is to be challenged then our criteria must be challenged as being inadequate. But, as we shall argue later, primarily in Chapter 5, our criteria give the best result we can hope for when seeking an ontic foundation for free will.

This brings us to a summary statement of what we intend for free will and what we do not intend:

First we take it that responsibility is important, indeed necessary, but we shall say little about moral responsibility. We say little because we take it that being responsible is a necessary precondition for being morally responsible; and because it is not clear, to us at least, what additional requirements allow the move from responsibility to moral responsibility. As we noted above, our models are most likely not equipped to facilitate the move.

Second, we shall also say little on the finer machinations of consciousness. We do not concern ourselves with dispositions, phenomenal states of consciousness, feelings and the like - we do seek to justify this approach later, primarily in the arguments presented in Chapter 5. But also it must be said, again, that our models do not look deep within minds but keep them solely as a facet of the universe and say little more.

Third, we do not worry to make the move from choosing to acting upon choosing. We take actions to be time-bound changes in the universe and that suffices for our needs. If there is any doubt about this stance then, whilst we stop our analysis at the moment of choosing, how things go on to become actions - in the sense of movements of arms, legs, mouths, speech and the like - is clear in our analysis.

And fourth, we do not seek a version of free will that hinges on the possibility of doing otherwise - we eventually argue, in Chapter 5, that such a notion is incoherent.[16]

[14]See Frankfurt (1969), Fischer (1995) and Fischer (2010).

[15]This was set in motion by Strawson (1962).

[16]Our arguments differ from those of Frankfurt (1969).

Now to what we do seek: we seek to produce a model of a universe that is inhabited by beings so that:

1. The universe presents genuine choices, genuine alternative futures, to its inhabitants;

2. The inhabitants 'choose' between these choices in such a manner that they are clearly and solely responsible for their choices, their selections.

We note that we use the term 'choice' in two different ways here and throughout this study: we use 'choices' to mean the alternative futures that are presented to some inhabitant; and we use 'choice' to mean the future that is selected by an inhabitant from the alternatives presented. The distinctions should be clear throughout but we shall often speak of 'alternative futures' for the former and 'selections' for the later when there is danger of ambiguity.

And we shall spend some time explaining and defining exactly what we mean by 'genuine choices' and what it is for some being to be 'responsible' for their choices.

We take **1** and **2** as necessary conditions for the exercise of free will. We make no attempt to argue that they are sufficient in themselves, however. There may well be much more to free will than **1** and **2** but we take it that **1** and **2** are complete as ontic conditions, albeit without explicit justification. Any additional conditions that might be argued for must then be conditions on the nature of the self in some way - requirements on the internal mental machinations of a self. So we might need to be able to believe we have free will, to be able to form intentions, to feel we have real choices and to feel we make genuine selections between them if we are to become morally responsible agents, or even just to possess free will in some guise. These issues play no part in what follows.[17] What we do hold to is that there must be necessary ontic condition if we are to have free will but, as we have said, we make no comment on whether ontic conditions exhaust the sufficiency requirements.

We do not make any claim that our final model is a faithful reflection of the universe we actually inhabit. In fact we argue that we cannot know if it is or not in section 5.13. Rather we seek to show there *is* a plausible model that will allow both these requirements to be achieved. What we shall make clear, and is hinted at in our discussion above, is that **1** requires the universe to be indeterministic and **2** requires a form of local determinism. As such we shall mould our solution within a model that has local deterministic and global indeterministic aspects - but we shall see that our solution is not Valerian. It is not Valerian as it does not attempt to inject indeterminism into the process of choosing, into the self's mechanism of selecting between choices.

[17] As such, we do not engage with capacity theorists and their arguments; see discussions in Dennett (1985), Frankfurt (1988) and Watson (2003).

The solution we offer, then, is one that rejects the possibility of achieving **1** and **2** in a purely deterministic universe. We do argue they can only be achieved in an indeterministic universe. But we shall see that we require a form of indeterminism that is not merely the simple opposite of determinism; rather, we require an indeterminism with specific facets hidden in amongst the universe's fabric. As such, our solution might be described as incompatibilist in that we reject the compatibility of determinism with **1** and **2**.

The aim of producing a model where **1** and **2** have the potential to thrive is what drives and justifies out foray into metaphysical models and our careful analysis of determinism, causality and supervenience. We look at determinism and indeterminism to set out their exact nature and scope. And as we do so we make it clear what assumptions about our universe we shall carry with us. We look at causality as we take it that it is the local underpinning of determinism and indeterminism - we hold that the way various bits of the universe, the objects as it were, interact is what is important about determinism and indeterminism. And then we define and examine supervenience. Supervenience is important for two reasons: first, it is often held the mental is in some sense distinct from the physical and supervenes upon it; and secondly supervenience offers the possibility of there being more than just causal relationships between various 'bits' floating about in the universe. If the mental is a separate realm that supervenes on the physical it may furnish routes to circumvent our dismissal of deliberation and the deliberations-to-choice-via-self-piloting. We must face this possibility. And if supervenience does show there may be ontic relationships between aspects of the universe, mental objects and physical objects perhaps, that are not causal we must show that these cannot have the power to slip another version of free will into the mix by bypassing the constraints of determinism and indeterminism that we argue for. In the end, to spoil the story, we show that supervenience does no harm to our arguments - unless, perhaps, we decide to accept an eccentric form of dualism; but even then it is not clear harm is inflicted.

Now we turn to outlining the arguments that we put in pursuing our aim.

1.2 An overview

In this section we set out an informal summary of all that follows. First a brief summary, then a detailed overview. All we say here is informal in that none of the technical definitions and mathematical machinery we construct is brought into play. Rather, our aim is to give an overview that glides over subtle details to provide, instead, a general orientation as to the arguments and ideas that we advance as we move towards our model for the exercise of free will.

As we orbit around free will we seek to build the foundations for an ontic metaphysics. By this we mean a metaphysics that admits to a real world of which we, as thinking beings, are some small part; and a metaphysics that seeks a consistent external perspective on the universe and ourselves. We do not advance arguments for this stance nor do we argue for its coherence or possibility as an approach. Rather we take it as a valid approach and leave it at that.

Our arguments begin at Chapter 2: there we draw the outer boundaries of the model for a universe that we shall be using. And in doing so we provide an ontic understanding of what it is to be a deterministic universe. Along the way we shall place some clear water between determinism and predictability. The key notion in Chapter 2 is that of a connexion-bound universe - roughly one that has a 'flow of influence' from moment to moment but with the details of what is mediating this influence, or how it flows, left deliberately vague.

Chapter 3 sets out a theory of causality for our connexion-bound universe and defines a number of concepts and ideas that will be central to the arguments and concepts laid out in subsequent chapters. In contrast to Chapter 2, which was about universes as wholes, Chapter 3 sets out to unpack the local machinations of the nitty-gritty bits-and-pieces that inhabit our universe. It tells us, in an abstract way, how specific stones break specific windows and the like.

Chapter 4 is an investigation of what it is for a universe to offer up genuine choices. The chapter is both a careful study of how ontic constraints affect a counterfactual approach to choices and also a preparation for the introduction of the idea of immanence. Immanence is our solution to the problem of choice.

Chapter 5 sets out to provide a model for the exercise of free will that satisfies the two criteria we set out above. It achieves this by first proving some technical theorems which set what constraints we must work under, and then by offering details of the model we claim satisfies our criteria. We then go on to argue that our model is as complete as is possible in the light of the constraints we outlined earlier in the chapter.

Chapter 6 investigates whether supervenience might affect our arguments. First, we set out a rigorous mathematical framework for supervenience. Then we argue that supervenience does not alter our arguments or conclusions, at least so long as we eschew eccentric dualist models of mind and world.

Chapter 7 concludes the work. Here we summarise the central argument we have presented. We point to weaknesses that we see as remaining and to where future work might fruitfully go.

Before we move to providing a more detailed summary of what is to follow, we must mention that we make much use of examples from theoretical physics throughout. We do so as physical theories provide convenient models that are both well-known and accessible in detail elsewhere. We also use physics because some of what follows is based on ideas and intuitions that come from examining the foundations and mathematical structures of quantum theory - even though no attempt is made to unpack quantum theory in the text or to lay out explicitly what these lurking motivations are. In saying all this, we should add a note of caution here: we do not take the models we use to be anything more than models and we remain fairly neutral on the finer points of the physics, at least here. Rather, the models serve as aides to thinking and understanding; they are used to illustrate ideas and to point suggestively to thoughts and ideas that underpin some of our notions. And it must be admitted that, on occasion, we use models that are not correct in as much as they do not reflect the modern orthodoxy - as far as it can be identified. For instance, we often talk of electrons as if they were billiard-ball-like entities whizzing around in space. This departure from the standard is generally harmless as, when we do stray from the accepted path, it is always clear that we have stepped away and we justify our wanderings by claiming the models we use are intuitively robust enough to illustrate our ideas.

Additionally, our models work on a naïve view of time. We ignore, for example, the complexities that special relativity might bring to our models. And, despite drawing on aspects of [non-relativistic] quantum theory throughout, our models ignore some of the difficulties that arise with some interpretations of quantum theory. For instance, the attribution of properties and the nature of objects in some interpretation could be seen to clash with aspects of what we say. We do, however, consider our models to be consistent with non-relativistic quantum theory, but at no point do we seek to explain or justify this claim.[18]

Now, a more detailed summary of what is to follow:

Chapter 2 is about universes. It is about the sort of universe we are going to take as basic as we move towards free will. Our universes evolve in time, changing from moment to moment. It is the way they change, the way each moment affects the next that sets the scene. There are two aspects to our model: one is that time is discrete so our universes evolve from discrete moment to discrete moment; the other is that there is something that allows each moment to affect or influence the next - or, perhaps, to flow into the next. Each moment in our universe is connected to the next moment and so influences it, or perhaps becomes it, through this connection. We call such universes 'connexion-bound'. The connexion might be a flow of energy-momentum, it might be a flow of information, it might be a careful mixture of the two or it might be something else altogether. We leave what flows, what allows each moment to become or influence the next, a little vague. If

[18]If pushed, we would favour a stance on quantum theory that admits both determinism and non-local causality.

we were pushed to give an intuitive feel for what we have lurking underneath the formal ideas, we would say that the influence is a flow of information carried through the form, or shape, of a medium; but we prefer not to make that explicit in the formal development as it is not clear that the terms we use - information, energy-momentum - are well-defined or even the correct concepts to use. But if it helps, as we mentioned, there is mostly no great harm in *thinking* that the form of each moment, the little details that make up the nitty-gritty bits of the universe, influences the form of the next moment through a flow of information or energy-momentum, or something similar.

We choose to use discrete time not because we endorse it as an idea - we remain largely neutral - but rather because it makes the discussion simpler and avoids a number of technical issues. If time is, in reality, continuous then we would hope that our models would be able to accommodate this continuity through some [mathematically based] limiting process.

To stand in contrast against our connexion-bound universes, we propose ‘Humean universes’. We name them after Hume although we make no claim they correspond to his ideas precisely. Humean universes are a sequence of moments - that is, universe states - one occurring after another in discrete time. But, in contrast to the connexion-bound universes, each moment in the Humean universe occurs without influence or transfer of anything from previous moments or to future moments. They stand alone and isolated as moments; the only relation they have to each other derives from their place in time.

We argue that a universe is deterministic if its present moment forces the next future moment to be exactly the way it is. That is, the present makes the future inevitable and it does so through the way the connexion between moments operates. Informally, the flow or influence from moment to moment must be the way it is and cannot be any different. We call the idea that the present state impels the future state to be the way it is ‘ontic necessitation.’ We say that a universe is deterministic if the current state always ontically necessitates future states. And if a universe is not deterministic it is then indeterministic. Humean universes are *all* indeterministic, whilst only those connexion-bound universes that do not, at every moment, ontically necessitate their futures are indeterministic.

What contrasts here with the usual approach is that we do not require the way [or, pattern in which] the universe ontically necessitates from moment to moment to be regular; that is, we do not require our deterministic universes to behave in roughly the same sort of way from moment to moment: they need not be predictable. So, for instance, we do not require billiard balls to move in straight lines, or planets to orbit in nice ellipses. Ontic necessity only requires one moment to impel the next to be the way it is; there is no requirement on the details of this necessitation. It is in this way that we start to separate out determinism from predictability. This is important as many arguments surrounding free will and determinism fall into using predictability as a manifestation of determinism; we reject that as mistaken.

In the last section of the chapter we distinguish between ontic and epistemic ‘laws of nature’. We hold that the confusions that plague understandings of determinism arise from - or lead to, depending on where one starts - the confusions that take laws as encapsulating the ontic.

Scattered throughout this first chapter are other issues. For instance, we claim that we cannot tell what sort of universe we actually inhabit - Humean, deterministic and so on; and we examine some arguments that suggest we can cite the patterns we perceive in the world - the regularity of the world’s behaviour - as evidence for our universe’s being connexion-bound rather than Humean. We conclude such arguments are invalid.

Next, we move, in Chapter 3, to unpick the nitty-gritty bits-and-pieces that inhabit our universes and to see how they relate from moment to moment. So whilst Chapter 2 examined only universes, Chapter 3 looks at billiard balls, electrons, people, cats, dogs, cheesecake and the like. We propose a model for causality which is somewhat different from the counterfactual models that haunt much of the literature. We propose that causality is a notion that captures the way bits of the universe evolve into, or influence, other bits of the universe later in time. And this influence is to be mediated through the connexion we proposed in Chapter 2.

We develop a formal, mathematically-couched, theory of causality by starting with relata as our basic building blocks. We do not take events, or anything else commonly proposed, to be relata. We prefer to leave our relata as primitive so as to give a scope of applicability to our model. If pushed, we might *think* of our fundamental relata as small pieces of space filled with energy and momentum; but this, it must be emphasized, is no more than a mode of thinking and not at all representative of the subtleties that lurk behind our notion of relata. We propose fundamental relata as indivisible causal units that exist at one moment of time and then influence fundamental relata at other future moments through the connexion. The advantage of this approach is that it allows us to develop a model without being side-tracked into arguments as to what constitutes an event - the usual causal unit - and even if an event, once defined, is really a causal unit. We just assume there is something at base that the universe is made from at each moment and that these ‘things’, our fundamental relata, enter into causal relationships across time mediated by a connexion.

Causality is then developed formally in terms of trans-temporal relationships between fundamental relata. In simple terms, a relatum at one time is part of the cause of a relatum at a later time if it can be shown that it has some influence on the form of the later relatum. And this influence is easy to trace through time as it is mediated by the connexion.

We set up the mathematical properties of fundamental relata in such a way that we can introduce a measure onto them. A measure is a way of quantifying something and here we use our measure to quantify both the causal content of a relatum and, by sleight-of-hand, to quantify the amount it contributes causally to later relata. We can, perhaps, *think* of the measure of a relatum as quantifying how much information it contains and, when used for causal influence, as a measure of how much information is transferred from one relatum to the other. Of course this is only one way of thinking and we do not push the information analogy at all in our formal definitions. This is partly because it is not clear that information is the apposite notion to use and partly because we do not wish to tie our model down more than we need.

The measure then becomes crucial to much of what follows in this chapter and later. In particular, we use it to propose a rigorous set of definitions for causal over-determination and causal epiphenomena as we shall have some need of both these ideas later.

We also examine, and dismiss, the technical idea of 'separability' - an idea we set as a contrast to essentialism. Essentialism is the notion that the causal [law-governed] nature of objects inheres in the objects themselves and separability is the idea that the causal nature of objects [or relata] is guided by separate laws that are set externally to the objects. For an essentialist, an electron's behaviour in a magnetic field, say, must be the same - same in obeying laws - no matter what universe and what time and place it encounters the field. Else, rejecting essentialism, the behaviour of the electron could change in other worlds if the laws of nature are changed there. We prove that our generous model of determinism must be accompanied by a broadly essentialist view of ontic laws. We also explain why our essentialist view differs in subtle ways from the conventional ideas of essentialism.

Chapter 4 marks the start of our work towards a model for the exercise of free will. We marshal the ideas we have developed in the previous two chapters to attack what it means for a universe to offer up genuine [ontic] choices. What we seek in this chapter is some ontic condition in the present that indicates the possibility of more than one potential future. And by 'ontic' we mean that we seek an actual aspect of the present that tells us of choices, alternative futures, yet to come and tells us in a manner that is not beholden to, or derived from, conceptual models held by intelligent inhabitants. We start our attempt to provide a solution to this by looking at counterfactuals - as they seem to encapsulate the notion of alternatives. We ask what makes it true, in simple terms, that something different from what actually happened at time t could have happened. So, again in rather over simplistic terms, we ask what makes it true that whilst the cue ball hit the red ball, it could instead have hit the yellow ball.

We start our investigation with a brief look at Lewis's possible world semantics in order to motivate our discussion. We reject it on two grounds: one, it relies on a broadly Humean/epistemic concept of the world and laws, something we have eschewed in Chapter 2; and two, because it does not take account of the nitty-gritty ontic machinations of the universe. We seek to resolve the second issue by setting out carefully how we might look at alternative possibilities - hitting the yellow ball instead of the red ball - if we are burdened with ontic constraints. We marshal the techniques and definitions of Chapter 3 to investigate the ontic constraints on counterfactuals in detail. We conclude that, in a deterministic universe, counterfactuals cannot launch and stay true to the ontic. Our arguments amount to pointing out that deterministic universes are made of a complex web of causally interconnected parts, and that what underpins the way these parts are forced to evolve, by obeying ontic laws, must be carried along in the fabric of the universe itself - so we take an essentialist stance for determinism. Any attempt to modify some small piece of a universe and keep its modification isolated is doomed to fail - the rest of the universe must fall in response to any modification. It is almost as if the modification is the fall of the first domino in a chain and there is no means to stop the cascade that must inevitably follow. We conclude that if we are to find a solution to our problem - namely what is the ontic nature of choices as alternative futures - we must look to indeterministic universes.

But indeterministic universes present us with a problem too. It is true that in such a universe there is no indication in the present as to where the future might unfold. I look with my ontic magnifying glass at all the little pieces that make up the universe now and yet find no indication at all as to what the future might hold or even if there is a future. And so it would seem that [pure] indeterminism does not provide an ontic underpinning for choices either. Rather, perhaps, we have been fooled into thinking it does as we can *conceive* of alternative futures lurking in the present of an indeterministic universe. But, we argue, these lurking futures are conceptual, they have no ontic underpinning.

We solve this difficulty by offering up 'immanence.' Immanence is the idea that certain types of indeterministic universes hold in their present some embryonic hints of alternative futures. And they hold these hints in the very fabric of the relata. These hints are not conceptual but are extant parts of the present. These immanent relata are what give truth to the claim that there are a genuine set of alternative futures lurking in the present; and so we may talk about them prior to their occurring and we may talk with confidence that our notions are ontically underpinned.

Chapter 5 is where we introduce our solution to the exercise of free will. Recall, we seek a solution that allows for the presentation of genuine choices to some being and for a choice, a selection,

then to be made in such a manner that responsibility is justifiably attributed to those that choose. We start the chapter with three theorems, using the techniques and ideas from previous chapters.
Our first theorem is essentially a restatement of an assumption we make. We assume that the mental - broadly construed - and the physical are both subsumed under the same single causal mechanism. We put this assumption as a theorem as we go on to argue that our assumption is *plausible* within the models we have set up in earlier chapters. Certainly there are different categories of experience: I experience pain, I experience the solidity of a table, I experience the presence of a table and so on. But it is less clear that a difference is experiential quality can be a justification for the idea that the physical, even in the face of mathematical physical theories, is causally distinct from the mental. Rather, we seek to argue, or make plausible, that all aspects of the universe - from thoughts, to the experience of pain, to billiard balls and the like - are subject to the same causal mechanism. We interpret Theorem 1 as giving us licence to talk of relata and their causal interactions without worrying as to how the relata cash out - we can rise above the issues of whether they are billiard balls, moments of pain, lumps of earth, electrons and so forth.

We then move to two more theorems: Theorems 2 and 3. These prove, in simple terms, that we as beings fully subsumed within the universe cannot self-pilot or pilot other relata to some specified end. They are demonstrations that it is not possible for any being or any object to actively seek to control where it will go or what it will do. We cannot control the full extent of ourselves; rather, we are creatures that must ride along with and within the universe, whether it be a deterministic or indeterministic one.

Theorems 2 and 3 show that we cannot bestow responsibility on any denizen of a universe by dint of their having chosen one course but having been able to do otherwise - in fact the notion of piloted deliberation, as we argued, collapses and cannot be made coherent. No being can control their deliberations in an active manner; no being can choose to go one way and then, on a rerun, choose to go the other: it is forbidden and our theorems show this. Piloted deliberation as a move from choices to choosing is dead in the water. As such, we argue that any notion of free will that invokes ideas such as ‘the ability to do otherwise’ must fail. There can be no such ability that is, in reality, any more than the result of mere random fluctuations. And so, if we are to seek what makes a being responsible for their [actions and] choices, we must look elsewhere. This leads us to our solution to the problem around which our work orbits.

We hold that a being is responsible for their actions if the being can be identified as the origin of its selections and those selections are consistent with the way the being is at the moment of choosing. And then we argue that to be consistent with the nature of a being can only be to be determined by the detailed structure of that being at the moment of choosing. So we claim that responsibility needs a notion of determinism - it may need more but we seek necessity not sufficiency. Selection from presented choices must be determined in full by the nature of the being doing the choosing. We carefully define a self, the being that is to be the mimic of a human, and find this self evolves deterministically: its nature at each moment is maximally determined by its nature at previous moments.

However, we still need there to be genuine choices, genuine alternative futures, for the world to offer up the possibility of turning left or turning right, of eating pie or eating cheesecake. We argue we need genuine choices because, in their absence and the rejection of pure indeterminism, all parts of the universe must work deterministically; and if the universe is deterministic our tentative definitions of responsibility and the self would give both a bullet and a psychopath equal cases to answer in any murder trial. We argue that local determinism is needed for responsibility and then we use our notion of immanence to provide the genuine choices that we also require.

Our model, then, has two facets: the determinism of the self and the indeterminism of immanence. We use the ideas of Chapter 2 to accommodate both within a single universe. The third strand of our solution is the combination of the two. We argue that to 'select' between genuine immanent choices and to be responsible for the selection can only be achieved if the result of the causal interaction between a self and an immanent relatum [alternative possibilities] is fully determined [by the self] at the moment they interact. And we show that we do need some form of counterfactual truth to be associated with free will [if the self is to be *solely* responsible for its selection] but our counterfactual is not one that involves alternative paths of deliberation.

We go on to argue, in section 5.13, that we can never know if we possess free will of our variety or not.

At this point we shall have achieved one of our stated aims. We shall have offered up a model that allows for the two necessary conditions we set out above for the operation of free will. Our model gives us both genuine choices and a notion of responsibility in the selection between those choices. It must be said that our solution cashes out choice and responsibility in a way that is not conventional; but it must also be said that the arguments we present show that there is no other route available if we accept the generously drawn outer boundaries of our models. As such, we put the case in Chapter 5 that our solution is the best that can be achieved.

At this point, we move sideways into Chapter 6 to examine supervenience in some detail. We do this both because there is a possibility that supervenience might provide additional ontic relationships that negate or challenge moves we make in some of our other arguments. Recall, our arguments assume that the only trans-temporal relationship is the causal one.

The first half of Chapter 6 is a careful, but brief, study of supervenience as it is found in the literature. We look only at sets of properties supervening on other sets of properties and ignore wider understandings - such as the supervenience of aesthetic qualities of art on the physical properties of paintings, or moral properties on aspects of actions, etc; but our models are general enough to be applied to other areas of consideration. We develop supervenience using ideas borrowed from elementary topology. In essence we show that any set of properties can be used to group objects together, turning the objects into points in a simple topology. And then we show that the supervenience relationship between the properties can be understood as a relationship between the corresponding induced topologies - that is, between the way the objects are grouped by each property set. As such we hold that supervenience of properties must be understood with respect to sets of objects. We develop some of the standard definitions that appear in the literature sufficiently to show how further developments could be accommodated using our mathematical framework.

In the second half of this chapter we seek to argue that supervenience does not affect our arguments. We explain how objects and properties are to be understood within the causal schemes we have used for our arguments. We then explain a number of different ways in which supervenience can arise within such causal schemes. We use these modes of supervenience to demonstrate that supervenience is not a relevant addition to our arguments in the sense that its presence does not change our conclusions in most cases of relevance. The dissenting case is one where supervenience is an ontic relationship between two [causally] distinct realms. We also suggest that a pure mode of supervenience, as we have defined it, arises from a second synchronous ontic relationship which we call, suggestively, ‘emergence’.

Chapter 6 concludes our arguments. In Chapter 7 we seek to review our models and arguments. We present a brief summary of our work, couched more technically than here, and we suggest where further arguments might be needed, where our models are lacking and what future developments might be interesting and fruitful.

Chapter 2

Determinism

2.1 Introduction

In this chapter we concentrate on defining the outer bounds of the models we intend to use for the tasks outlined in Chapter 1.

We have two aims here: one, to set out clearly what we are to mean by determinism and indeterminism, both ontically construed; and two, to set out clearly the assumptions adopted for the models we shall be using in this study. As such, this chapter concerns universes largely as wholes and we postpone unpacking their internal machinations, which we subsume under the idea of causality, to the following chapter.

2.2 Introducing determinism

In simple terms, determinism is the doctrine that the future is forced to be the way it is by the present and by the way the universe evolves. This rough conception is usually unpacked by paraphrasing the standard way of encapsulating determinism, namely that found in Laplace. He says,

> We ought then to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instance an intelligence which could comprehend all the forces by which nature is animated and the respective situations of the beings who compose it - an intelligence sufficiently vast to submit these data to analysis - it would embrace in the same formula the movements of the greatest bodies in the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes. (Laplace, 1902, p. 4)

We can see here that Laplace couches his idea in heavily epistemic language. He can be thought of as characterizing determinism as follows: if we could wander around the universe and make a complete inventory of every part of it to the most exacting detail then, if the universe is deterministic, this inventory together with the laws of nature would be sufficient for us to calculate the exacting details of future states of the universe. Of course we would also have to have access to these laws and it must be a condition of Laplacian determinism that the universe is governed precisely and always by these laws.[1] The practicalities of Laplace's idea are clearly beyond us, but as we can conceive of an extremely intelligent all knowing being, or perhaps conceive of making ourselves all powerful, then the idea seems plausible.

There are a couple of aspects to Laplace's characterization we should note. First, if we are inhabitants then our trudging about disturbs the surroundings and, additionally, it takes time to wander here and there examining the state of things. Further, we need to be able to observe and record ourselves too in the inventory as our presence will form part of the universe. Clearly, taking Laplace too literally furnishes difficulties. To mollify these difficulties we must place Laplace's being outside of the universe-proper and grant him the power to peer in and absorb the full exacting details of the universe at any instant. Second, it is usually taken that the laws of nature are fixed throughout time, so that the same basic principles and schema of calculations can be applied at any time by our Laplacian super-being. And such laws of nature are usually taken to be something akin to a finite mathematically-couched set of rules which, when applied to the data that represent the current state of the universe, give a complete description of future states. And so we must assume that the Laplacian being is able to interpret the relationship between the mathematics, or his calculation schema, and the data he gathers from the universe with ease. Third, Laplace mentions cause and effect. We deal with the relationship between ontic determinism and causality in section 3.8 where we shall explain that causality alone is not sufficient for determinism to obtain.

Laplace's position, as we have it, is that for a universe to be deterministic is for it to be perfectly predictable using some systematic summary of the laws of nature and a complete inventory of the world at any time. This approach to determinism is frequently summarised in one form or the other in the literature. For instance, Roskies (2010) holds that determinism is . . .

> . . . the thesis that the state of the world at a given time plus the laws of nature uniquely determine the state of the world at every other time. In other words, given an exhaustive specification of the physical state of the universe at any given time, as well as perfect knowledge of the physical laws, one could, in principle, predict with certainty the physical state of the universe at any other time. (Roskies, 2010, p. 154)

[1] We leave the idea of 'laws' a little vague here. We tackle them in more detail in section 2.10.

In this chapter we are going to advance a view of determinism which keeps broadly to the basic spirit of Laplace but differs in detail substantially. We are going to take it that determinism is a doctrine about the universe itself, an ontic doctrine, and not about the way we describe the universe.[2] Recall, in Chapter 1, the ontic approach was characterized as having a number of facets: there is a real world external to us and we are justified in talking of this world as if in a position from outside looking in - we shall characterize this type of talk as coming from an Archimedean perspective in section 2.5;[3] additionally an ontic doctrine is one that concerns things as they are in themselves and not the way we might choose, or be forced, to describe them or interact with them. In contrast, an epistemic perspective is one that concerns our knowledge of the world gained from what access we have to it, and then also concerns the ways we break down and describe this knowledge - we call this, broadly, talk from an internal perspective in 2.5 below.

We shall hold that a universe is deterministic if and only if the future is made necessary or inevitable in some manner by both the present universe **and** by the 'nature' of the universe. We shall set out the details of this below by moving to a universe that is connected together from moment to moment. To make this idea of connected parts clear we contrast it throughout with its antithesis: an idealised Humean universe - something we define carefully below.

Whilst we only allow universes themselves, as extant things, to be deterministic we should note that the idea of determinism is sometimes associated closely with scientific theories instead.[4] For instance, Popper (1950, p. 120) states that he intends to take "...the word 'determined' as 'predictable in accordance with the methods of science.'"[5]

There are careful studies that look in detail at how certain views on determinism and current physical theories fit together: the *locus classicus* is Earman (1986).

2.3 Defining determinism

To launch our discussion, we propose an informal simplified definition of determinism and then progress toward a more rigorous definition by unpacking our terms:

> **DEFINITION 2.1: Deterministic Universe [informal]**
>
> A universe is **deterministic** if its current state and the laws of nature make future states necessary.

[2] An interesting perspective on an ontic construal of determinism in the context of quantum theory can be found in Atmanspacher & Bishop (2002, pp. 49-74).

[3] We must be careful not to take this analogy too far. Laplace's demon does not merely peer in from the outside, as we might peer into a box from the outside, but rather the demon has access to the internal machinations of all aspects of the universe in a manner that does not merely restrict him to a 'Mary's' external viewpoint - see Jackson (1986).

[4] See Butterfield (2005, p. 195) for a brief discussion.

[5] Popper is not immune to difficulty when dealing with determinism: see Earman (1986, pp. 8-10).

Here we can think of the 'current state' as the full contents of a time slice of the universe - a single moment in a universe's evolution. Below we shall explain how this fits with our assumptions on time for our model.

Two aspects of our tentative definition need some clarification: what it is to 'make future states necessary' and what we are to mean by 'laws of nature'.

We are going to leave the precise characterization of a universe until section 3.7; here we take it as a primitive term and outline its general characteristics. In section 3.7 we shall characterize a universe, using the connexions between its states at differing times, as being a closed causal nexus. For the moment it is sufficient to *think* of this connexion-bound universe as being constructed from everything that impacts on us in some manner or that we impact on, and also those things that impact on those things that impact on us, or are impacted on by those things ..., or impacts on those things that ... and so forth iteratively.

We shall now turn to outlining some assumptions and constraints we shall carry with us in our models of universes.

For simplicity we take a slightly naïve view of time. We assume that the universe evolves in time and is minimally constituted of a linear sequence of well-defined **moments**, or **states**. The requirement for evolution in time is adopted to avoid unnecessary complications that might arise were we to allow block-universe scenarios into the mix.[6] And, further, we ignore issues related to special relativity throughout this study.[7]

In order to be able to deal cleanly with Humean universes and contrast them with connexion-bound ones, we place some requirements on time for our models.[8] Each moment is required to be well-defined so that there is no ambiguity as to what exactly is included within its scope; and moments are assumed to form a linear sequence so that there is no doubt as to their time ordering. We also require that there is a clear meaning to the idea of 'the next moment' following any given moment. This last assumption requires that time must be discrete and, depending on the life of the universe, the set of moments is either finite or countably infinite. We also largely ignore issues as to the extent that time stretches backwards - that is to say whether it has a 'beginning'. Taking these assumptions allows us to use numerical subscripts to indicate temporally juxtaposed moments.

Our time labelling convention runs as follows: we shall denote the current state of the universe as $\mathfrak{U}(t_0)$ and the state immediately following as $\mathfrak{U}(t_1)$ and immediately preceding as $\mathfrak{U}(t_{-1})$. When we want merely to indicate earlier and later states without being so specific as to how close they

[6] A discussion of the philosophical and physical ideas behind block-universes can be found in Price (1997).

[7] A clear mathematical summary of special relativity can be found in Schutz (1985, Chapter 1), and a more philosophically based account in Bohm (1996).

[8] We make some further comments on the requirements for time we adopt here in section 2.4 below.

are temporally, we shall use letters as subscripts and time ordering will then correspond to alphabetical order. So $\mathfrak{U}(t_c)$ is earlier than $\mathfrak{U}(t_f)$ but later than $\mathfrak{U}(t_b)$. We use this convention for time ordering throughout this work.

In making these assumptions we are avoiding a number of difficult issues that arise from adopting a continuous approach to time and implicitly assuming that they do not affect our arguments. For instance, if we were to take time as continuous, or even as just parametrized by a dense set on the real line,[9] we would have no well-defined way of discussing the state immediately following or preceding some other state: this is why we chose to use the integers, $\mathbb{Z}$, to parametrize our moments rather than the rationals, $\mathbb{Q}$ - the former gives us a clear notion of 'next' and 'previous' whilst the later does not, despite both sets having the same cardinality.[10] Additionally we note discrete time is a possibility entertained by some current physical theories.[11] What is less clear, if we take time to be discrete, is how we are to understand our epistemic access to the universe - that is the way the universe impinges on us and furnishes us with knowledge of its make up. We shall say more on this below when we discuss how we are to use the idea of necessary connexions in characterizing determinism.

We return to issues involving discrete time in section 2.4 below where we briefly summarise a number of issues and difficulties that arise in connection with our assumption that time is discrete.

Now we have set out some of the basic assumptions of our model, we turn to characterizing the nature of the necessity that makes future states inevitable on the back of current states. We base our conception around the idea of a **connexion**. A connexion is our name for the ontic - i.e. real - influence or transfer that we take to occur between juxtaposed states. It is via the connexion between states that one moment can come to shape or influence another. And we shall use terms such as 'influence' and 'evolve into' freely when talking of these connexions. Eventually, in Chapter 3, we shall link together notions of causality and connexions so that to be caused, in simple terms, is to be effected via a connexion.

There will be talk, then, of connexion-bound universes. And these shall be contrasted with Humean universes. Humean universes, as we shall explain, are to be ones composed of sequences of moments, but sequences unencumbered by a connexion. We use the contrast to help emphasize and explore aspects of the connexion-bound structures.

Now, we turn briefly to Hume to launch an outline of Humean universes. On standard readings, Hume (2008) claims that we derive our notions of causality from our observations of constant con-

[9] See Bourbaki (1998) for definitions and a rigorous treatment of dense sets.

[10] It is a characteristic of continuity and dense sets that for any two times, or points, there are always other times, or points, between them.

[11] See for instance Yang (1947), Rovelli & Smolin (1995) and Bojowald (2001).

junctions and that there is no more to causality than this.[12] Hume's conception of cause does not allow for any form of connection, and indeed connexion, between one moment and the next; that is, one moment has no impact or influence on any other moment that follows it: there is no transfer of energy, information or anything else. For a Humean universe there is just 'one damn thing after another'[13] but with nothing binding them together. The fact that the world about us appears to be 'regular' and to obey 'rules' is, in this Humean conception, the result of pure serendipity. Further, there is no instantaneous relationship between objects, or things, at one location with those at any other within each moment. This Humean outlook is best summarised by Lewis (1987).[14] He says,

> Humean supervenience is named in honor of the greater denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. (But it is no part of the thesis that these local matters are mental.) We have geometry: a system of external relations of spatio-temporal distance between points. Maybe points of space-time itself, maybe point-sized bits of matter or aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that. (Lewis, 1987, p. ix-x)

Here, if we were to admit this Humean universe, then it would be exhausted by the sequence of moments. There would be nothing ontic, nothing actual, connecting one moment to another. A Humean universe, for us, is constituted solely of a sequence of well-defined moments, or states, parametrized by discrete time; and each state is constituted of local configurations of objects, whatever 'objects' might be.[15] We do not here argue for or against this view: our Humean universe is to act as a foil against which to test our concepts and arguments. Were we to go down the path of unpacking a Humean universe in full detail, we would need to specify the internal structure of each moment, gain some understanding of Humean-objects and their persistence, explain how we gain access to these objects as denizens, and so forth. Instead, we are going to take a different approach and, as we progress, investigate the ramifications of adding to the basic Humean universe

---

[12]This may well be a naïve 'standard reading': see Beebee (2011, p. 1). An alternative perspective can be found in Strawson (1989) and a critique of his approach in Beebee (2006).

[13]The phrase seems to be attributed to a number of different people.

[14]It is worth noting in passing that much of the sentiment expressed by Lewis in his introduction to Lewis (1987) stands against some of the empirical and philosophical aspects of quantum theory. Lewis is aware of this and dismisses this as a worry. On page xi he says, '...I am not ready to take lessons in ontology from quantum physics as it now is.' Here, perhaps in contrast to Lewis, we *are* motivated by the issues and counter-intuitive ideas that are scattered throughout quantum theory and, indeed, physics in general.

[15]It does seem that to maintain such a view we must have space and time ready-made as receptacles in which to place the content of each moment and the sequence of moments.

extra structure beyond our minimal sequence of well-defined states and of making a few additional assumptions concerning the internal structure of each moment. These bits of extra structure are only to help us contrast the connexion-bound universe we develop with ones where such connexions are absent.

So, to recap: for the universes that will be central to this study, we require there to be an ontic connexion between adjacent moments; recall, that means we require each moment to be influenced by, in some broad manner, the moment previous to it. This is an actual influence, an ontic connection across moments. We make no claims as to how this influence is propagated. It may be through a transfer of energy-momentum or information or matter or something altogether different. The only point is that, in contrast to a Humean universe, we have this extra aspect. We could, of course, postulate connexions between non-adjacent moments, or one that is backwards looking in some manner, but we do not do so: we take the view that adding such complications provide more egging of our pudding than is required.

Using this connexion we can now explain the way in which the future is necessitated on the back of the present in a deterministic universe.

Recall, a connexion between [temporally juxtaposed] states means that the presence of one state has some influence or impact on the other, later, state. If this influence results in the future state's inevitability then we have it that the future state is necessitated by the current state. We shall say that, under these circumstances, the future state is **ontically necessitated** by the current state. So, for the determinism advocated here, the current state ontically necessitates its immediate future state via the connexion between them. And by a temporal concatenation of such necessitations, any state ontically necessitates any state future to it in a deterministic universe.

When we come to provide a rigorous definition of determinism it will be important to recall that the presence of a connexion is itself not sufficient to ensure the future states are ontically necessitated by the current state. It could be the case that the connexion is such that it does not make all, or any, of the future states inevitable. For instance, it may only partially constrain some future state in some manner. And later, here and in Chapter 3, we shall see that connexion-bound universes need not be deterministic. Additionally the current state, in a deterministic universe, can only be considered sufficient for the future state. It may be the case, for instance if there is some over-determination or epiphenomena,[16] that non-unique subsets of the current state are themselves sufficient to make future states ontically necessary. Or it may be the case that one single sub-state

[16]These are defined rigorously in section 3.6.

of a current state is all that is required. We investigate the interplay between the connexion and the states in more detail in Chapter 3 when we introduce causal notions into the discussion.

We shall indicate ontic necessity relationship using the symbol $\boxdot\!\!\rightarrow$. For instance, to summarise the statements that a complete state of the universe at time t_0, denoted $\mathfrak{U}(t_0)$, ontically necessitates the state $\mathfrak{U}(t_1)$ we can write:

$$\mathfrak{U}(t_0) \boxdot\!\!\rightarrow \mathfrak{U}(t_1) \tag{2.1}$$

sometimes calling $\mathfrak{U}(t_0)$ the antecedent and $\mathfrak{U}(t_1)$ the consequent of this ontic necessitation.
Throughout this chapter, when we talk of the state of the universe we shall always mean the complete state - those unencumbered by issues of epiphenomena or over-determination and such like. It is only when we look at causality in Chapter 3 that we can start to move the discussions of determinism to parts of universes.
Of course, it is important to remember that (2.1) is a symbolic summary of an actual, or ontic, relationship in the universe. The use of $\boxdot\!\!\rightarrow$ is always taken to be accompanied by the implicit assumption that there is a connexion between the antecedent and the consequent.
We can also extend the use of this notation by requiring that, for a deterministic universe, ontic necessitation must be transitive across states. It does not seem unreasonable to take it that if $\mathfrak{U}(t_0)$ ontically necessitates $\mathfrak{U}(t_1)$ and then $\mathfrak{U}(t_1)$ ontically necessitates $\mathfrak{U}(t_2)$ then it is also the case that $\mathfrak{U}(t_0)$ ontically necessitates $\mathfrak{U}(t_2)$. So we can extend the use of our symbol to any countable concatenation as follows:

$$\mathfrak{U}(t_a) \boxdot\!\!\rightarrow \mathfrak{U}(t_z) \tag{2.2}$$

by considering it as shorthand for the sequence:

$$\mathfrak{U}(t_a) \boxdot\!\!\rightarrow \ldots\ \mathfrak{U}(t_m) \ldots \boxdot\!\!\rightarrow \mathfrak{U}(t_z) \tag{2.3}$$

And, as mentioned, in such cases the connexion between the antecedent and the final consequent is really a series of intermediate connexions from moment to temporally-juxtaposed moment.
We must be careful with ontic necessity as, contrary to many conceptions of determinism, it need not be time reversible. That is, if one state ontically necessitates the subsequent state, making it uniquely necessary, there is no guarantee that the later state, taken in isolation, must then neces-

sarily have arisen from the earlier state.[17] It is quite possible that it, as an isolated state, could have arisen from one amongst many possible earlier states. In simple terms, ontic necessity need not be time reversible and so Laplace's claim that both the future and the past are deducible from the present state need not be a feature of it. If we allow ontic necessity to work in both temporal directions, both forwards and backwards, then we can write:

$$\mathfrak{U}(t_0) \boxdot\!\!\rightarrow \mathfrak{U}(t_1) \Leftrightarrow \mathfrak{U}(t_1) \boxdot\!\!\rightarrow \mathfrak{U}(t_0) \tag{2.4}$$

and by extending (2.2), we also have:

$$\mathfrak{U}(t_a) \boxdot\!\!\rightarrow \mathfrak{U}(t_z) \Leftrightarrow \mathfrak{U}(t_z) \boxdot\!\!\rightarrow \mathfrak{U}(t_a) \tag{2.5}$$

We shall return to this in a moment when we define retrodictable determinism.

We can now provide a definition of a type of determinism that will be adequate for our purposes here. It is perhaps noticeable at this stage that, in contrast to the sentiment expressed by Laplace above and **DEFINITION 2.1** in section 2.3, we have not yet mentioned the laws of nature. We shall examine their place in our models below, in section 2.10.

DEFINITION 2.2: **Connexion-bound Universe**

A **connexion-bound** universe, $\mathfrak{U}(T)$, or usually just **universe**, is a countable set of states ordered by a time parameter in which temporally juxtaposed states are related through a connexion. We write: $\mathfrak{U}(T) = \{\mathfrak{C}(t), \mathfrak{U}(t); t \in T\}$ where T is the set of all times.

Here we use $\mathfrak{U}(t)$ to represent states of the universe and we use $\mathfrak{C}(t)$ to *indicate* the presence of a connexion-based relationships between the states.

This allows us to construct a set of definitions:

DEFINITION 2.3: **Deterministic Universe**

A connexion-bound universe, $\mathfrak{U}(T)$, is **deterministic** if and only if $\forall \mathfrak{U}(t_n) \in \mathfrak{U}(T)$ it is the case that $\mathfrak{U}(t_n) \boxdot\!\!\rightarrow \mathfrak{U}(t_{n+1})$.

Recall, we set a convention that general temporally ordered states are labelled by letters and the alphabetic order of the letters reflects the temporal ordering of the states. So, for any deterministic

[17] Here we mean to say, using the language of Chapter 3 for a moment, that two states are equal if they are constituted of identical configurations of relata and that the relata in the same places in each of the equal states are themselves content-identical. For definition of 'content-identical', see **DEFINITION 3.13** in section 3.9.

connexion-bound universe, we can also write:

$$\forall \mathfrak{U}(t_a), \mathfrak{U}(t_b) \in \mathfrak{U}(T) : \mathfrak{U}(t_a) \boxdot\!\!\rightarrow \mathfrak{U}(t_b) \tag{2.6}$$

We can go further and define the idea of retrodictability as follows:

DEFINITION 2.4: **Retrodictable Universe**

A connexion-bound universe, $\mathfrak{U}(T)$, is **retrodictable** if and only if it is both deterministic and $\forall \mathfrak{U}(t_{n+1}) \in \mathfrak{U}(T)$ it is the case that $\mathfrak{U}(t_{n+1}) \boxdot\!\!\rightarrow \mathfrak{U}(t_n)$

and we can then write, by iteration of this definition as before, that for retrodictable universes:

$$\forall \mathfrak{U}(t_a), \mathfrak{U}(t_b) \in \mathfrak{U}(T) : \mathfrak{U}(t_b) \boxdot\!\!\rightarrow \mathfrak{U}(t_a) \tag{2.7}$$

We have only given definitions of determinism for connexion-bound universe. What about Humean universes? The answer is that they are *not* to be considered deterministic as there is no mechanism for one state of the universe to influence another - recall that a lack of influence was a fundamental feature of Humean universe.

We define a Humean universe as follows;

DEFINITION 2.5: **Humean Universe**

A **Humean universe**, $\mathfrak{H}(T)$, is a countable set of states ordered by a time parameter in which states are not related to each other in any way other than time ordering. We write: $\mathfrak{H}(T) = \{\mathfrak{H}(t); t \in T\}$ where T is the set of all times. Further, there is no other relationship, beyond spatial juxtaposition, between the ontic-inhabitants that exist within each state.

Where by 'ontic-inhabitants' we mean objects, constituents of the universe and the like but loosely construed.[18]

And, as there is no connexion, there can be no ontic necessity operating. It is for this reason that we hold that a Humean universe is not deterministic.

This allows us to set out exactly what we mean by an indeterministic universe:

DEFINITION 2.6: Indeterministic Universe

A universe is **indeterministic** if it is either a Humean universe or it is a connexion-bound universe which is not deterministic.

[18]We say 'loosely construed' here as we remain neutral as to what the constituents might be. Additionally, there are difficult issues concerning the unpacking of juxtaposition in a Humean universe - unless we choose to assume an underlying spatial manifold on which the play of objects occurs. We ignore these issues here as they are not central to our concerns and because we are taking a Humean universe to be a model for contrast alone.

And finally noting that it is possible for an indeterministic universe to have deterministic features - that is, it may well be the case that *some* of its states necessitate following states but determinism cannot be claimed unless they *all* do.

2.4 Issues with assuming a discrete time

There are some issues and difficulties with our model that arise from using discrete time parametrized by integers that we shall discuss briefly. We mention them here because discrete time is an important feature of our model and we wish to make clear both what we are assuming for our models and what issues we are putting aside. Recall, we introduced discrete time in order both to assist with setting out our model of a Humean universe and to ensure we have a clear idea of temporal state juxtaposition - so we can talk about 'the next state' and 'the previous state' and so forth. Here we briefly review the main difficulties one at a time:

1. Temporal Gaps: We have labelled our times using integers. We have taken the current moment to be labelled using 0 and the moment following it by 1 and the moment preceding it by -1. As such our labelling convention allows time to stretch back and forward in time in an infinite and countable set of moments. However, we must be careful not to take the properties of the parameter set across to being properties of time. For instance, when we think of the integers we think that there are gaps between them filled with other numbers. We think this way because we view the integers as being embedded in a bigger set of numbers that occupy these gaps. So, for instance, there is a gap between 0 and 1 filled with an uncountably infinite set of numbers. It is easy, but wrong, to then move this analogy across to moments and talk, for instance, of the gap between t_0 and t_1. There is no temporal gap between these two moments. We must not think of moments as being separated in any manner by temporal gaps; rather the moment at t_0 is next to that at t_1: there is no temporal gap between them.

2. Duration: If we take two times, 12pm and 1pm say, then we know there is an hour of time between them. This is because we assume, in this case, that time is continuous and fills the gap between 12pm and 1pm fully. That is we have a sense of duration that derives from the assumption that time is continuous. However there is a problem with duration and discrete time. In simple terms we can set this out as follows: each labelled moment in discrete time has no duration - it is infinitely thin so that its temporal width is zero. If we make of pile of these infinitely thin moments it might seem that they will eventually produce a pile of moments with a height - somewhat like a pile of playing cards perhaps. However, this is not what happens; our analogy with playing cards is wrong as playing cards have a small, finite thickness that build up to give the pile a height. In contrast our pile of moments has no thickness to build up: in fact, a countable pile of infinitely thin moments must have a height of zero.[19] We know our pile of moments must be countable as it is labelled by a countable set and so if we adopt a notion of discrete time we lose any notion of temporal duration

3. Moving from discrete to continuous: We would hope that our model using discrete time would act as an approximation for a model that uses continuous time. That is, there must be some rigorous method of turning our discrete time models into continuous time models without losing the structures and ideas we associate with our model - such as causality in Chapter 3. Mathematically there are ways of achieving a transition from a countable infinite set to a continuum.[20] But the worry is whether mathematical methods concerning the indexing of moments can be argued to carry with them ontological possibilities and to preserve the features of models based on discrete time. It is not clear, especially when considering points 1 and 2 above, that this is possible and, if it is, how it could be achieved.

4. Persisting objects: Each moment in discrete time is self contained and distinct from previous and subsequent moments. If this is the case we must ask what we are to mean by an object that persists across time and how we might understand what persistence across time might mean.[21] In this study we achieve our idea of persisting object through a notion of causality and connexions but we leave detailed worries as to object-persistence to one side.

5. Dynamic features: When we observe the world we perceive some continuity from moment to moment. We have some notion of what it means to travel with some speed and to move about. However, discrete models of time lack notions that derive from assuming a continuous underlying time. Our models offer at best a concatenation of static frames. We might wish to consider each state in a universe's evolution as a dynamic state, possessing instantaneous velocities and the like. If we do want these dynamic aspects then we must admit it is not clear what an instantaneous velocity, for instance, would be in such a case. This is because instantaneous velocities are defined mathematically as limits and so involve continuity.[22]

[19]This can be justified formally using measure theory. See Halmos (1974).

[20]For instance, we can transit from the integers, $\mathbb{Z}$, to the rationals, $\mathbb{Q}$, by the addition of multiplicative inverses. We can then transit from the rationals to the real numbers, $\mathbb{R}$, by requiring every Cauchy sequence to converge. A sophisticated treatment can be found in Garling (2013).

[21]We leave the terminology loose here. Our worry is how we might identify what aspects at each moment are to be counted as important for questions of persistence. There is a sophisticated debate about objects' 'endurance' and 'perdurance' that we do not wish to enter here. See Sider (1999), Sider (2003) and Lewis (2001b).

[22]Definitions can be found in any elementary book on calculus; see, for instance, Spivak (2006).

These issues amount to the fact that adopting discrete time is not an innocuous and simple move. As our thinking tends to be in term of continuous time it is easy to take this thinking across uncritically to our understanding of discrete time and this can lead to misunderstandings or false arguments. In this study we place time firmly in the background. We chose to take a naïve and simple view of time as discrete and to put aside the issues we have outlined above. We shall talk as if the analogies between discrete and continuous time are unproblematic and we make the assumption that our models, based on discrete time, will remain robust under a rigorous transition to continuous time. Our decision to ignore these, and other,[23] issues with time is certainly a lacuna of this study and leaves many questions for future clarification.

2.5 The Archimedean constraint

In the following two sections, 2.6 and 2.7, we look at some issues arising from applying a natural metaphysical constraint to the Humean universe and contrasting it with our connexion-bound model. This will allow us to explain some further assumptions we must make for our connexion-bound universe to ensure it is a robust model. In this section we set out the metaphysical constraint that will motivate our arguments in sections 2.6 and 2.7.

There are two broad perspectives on the universe - one from without and one from within - that are frequently used implicitly in philosophical discussion that accepts realism in some form or other. Drawing an explicit distinction between them is useful for some of the discussions that follow.
Often we talk of the universe and its machinations as if from a viewpoint entirely outside and separate from it. That is a viewpoint where we look in at the world from without, and in looking we apprehend all that is within but in a manner that does not affect it in any way whatsoever.[24] We call our external viewpoint - the one that is Laplace's demon's view - the **Archimedean viewpoint**. For instance, much of what has gone before in this chapter is an attempt to view and deconstruct the universe from an Archimedean perspective.
The second viewpoint is one we have from within - from a perspective of being full inhabitants of the universe and constrained by our being fully within it. We call this the **internal viewpoint**.

[23]For instance we ignore issues with time that might arise from quantum theory - for instance the energy-time uncertainty relation - or other areas of physics such as special and general relativity. For instance, see discussions in Landsberg (1984), Sklar (1992) and Price (1997).

[24]And we take it that this viewpoint is complete in that it is not merely a 'Mary's view' on things - see Jackson (1986). That is, we must not take the idea of an external view of the sort we might have - as beings *within* a universe - when looking into, say, a cardboard box as being analogous to the external viewpoint of some 'being' looking into the universe from a position outside of that universe. The latter view, from outside of the universe, is the one we attribute to Laplace's demon. We discuss this further in section 5.5.

If we take a metaphysics based on a purported Archimedean perspective it will be a model with a basic set of assumptions, which we call **axioms**. Additionally, in many cases the internal viewpoint is itself subsumed under the metaphysics and so beholden to these axioms and related structure. If the way the internal structure is approached extends, or potentially extends, beyond the encompassing metaphysics in any manner then the internal perspective will either stand in contradiction to the metaphysics under which it is subsumed or else require further assumptions to be added to the axioms of its enveloping metaphysics.

So we hold that any metaphysical scheme that purports to model the world and its inhabitants must obey a constraint. The constraint says:

DEFINITION 2.7: The Archimedean Constraint

Given any metaphysical axiomatic system concerning the universe from an Archimedean perspective, then any assumptions that govern the internal perspective of its inhabitants and the machinations of objects within the universe - all of which are subsumed under the metaphysics - must either be consistent with the encompassing metaphysical system or be capable of being made consistent by the addition of further axioms to the system.

We offer examples of the application of the Archimedean constraint in sections 2.6 and 2.7 below.[25]

Informally, this requires that everything must fit together nicely in any metaphysical theory: we must be wary of the assumptions we make when we analyse how we experience the world once we accept our metaphysical framework as setting the scope of our discussion. Our talk must not overstep any bounds or add new structure that stands in excessive tension with what is already proposed. And we must not take talk that is primarily internal to be talk that is Archimedean or vice versa. The Archimedean Constraint is, then, here an axiom of meta-metaphysics. We leave what it is to fit together nicely, to lack internal tension, to be non-contradictory etc, as primitive notions.

[25] As an example: it might be argued that the 'measurement problem' in some approaches to quantum theory is an issue that exhibits a tension. In the measurement problem there is a tension between the axiomatic structure - quantum theory - and aspects that should be subsumed under it - namely, measuring devices and their behaviour. That is, the axioms of the theory suggest all equipment in the universe must be governed by the axioms; but still there is, when discussing measurements, a tacit requirement that measuring apparatus acts classically and not quantum mechanically. Discussions of the measurement problem can be found in Isham (1995, pp. 175-188) and Bohm & Hiley (1993). For a philosophically-couched discussion of measurement in quantum theory, see Jammer (1974, Ch 11).

2.6 The problem of regularity

Here we shall use the Archimedean constraint to examine our Humean universe model. We do so to understand what the claim that we inhabit a connexion-bound universe entails and to set the ground for the introduction of a further axiom that we shall need to add to our connexion-bound model. The extra axiom concerns time and is discussed in section 2.7 below.

The usual arguments hold that a Humean universe is less likely to be a robust model of our universe because our universe appears to us to be 'regular'.[26] That is, the epistemic access to the world we have has certain features, which we take to be reflective of the actual world. And then because the view we gain from this epistemic access is fairly amenable to theory - it changes continuously, and is generally both well-behaved and consistent - it is argued that this 'regularity', which is assumed reflective of an ontic regularity, must be derived from more than mere serendipity.[27]

In this, and the next section, we shall explain the move from regularity to a rejection of Humean models is in error. The usual argument to rejection attempts to claim that Humean models are not equipped to explain regularity as arising from anything other than fortuitous chance. Rather, here we propose that regularity alone is either neutral as a discriminator between universe types or, with the addition of some simple assumptions, we find regularity favours the Humean over the connexion-bound. If we then add the further requirement that we be able to explain how our universe impinges upon us, as it must if we are to perceive it, then we find we can argue that a connexion-bound universe provides such an explanation whilst the Humean cannot. However, in arguing for a connexion-bound universe on the grounds that it has the axioms to explain our epistemic access to the external world, we find regularity raises its head to cause some trouble. This last issue we tackle in section 2.7. It is there we find we must add a further axiom to our connexion-bound universe.

We now turn to flesh out these arguments in stages. We start by looking at the Humean case and explain why we hold that regularity alone does not allow the move to asserting the presence of connexions. Let us summarise the axioms we require for Humean universes:[28]

> **[O-SEQ] ORDERED SEQUENCES**:
>
> The states of the universe form an ordered sequence [parametrized by a discrete time];

---

[26] For instance in Strawson (1989).

[27] We must caution that we use the term 'regular' in a slightly loose and intuitive manner: we assume regularity is compatible with a number of universe types - so that some uncertainty as to the future, perhaps as in some versions of quantum theory, would still count as being regular here.

[28] We also require two further axioms which play no part in our discussion: **[Aa]** The world [or universe] is real; **[Ab]** The content of the universe is well-defined at each time parameter value.

[NO-CONX] NO CONNEXION:

There is no connexion between one state and the next;

[ISOLATION] ISOLATED OBJECTS:

There is no internal ontic relationship [other than simple spatial juxtaposition] between the various objects populating each state.[29]

When we indulge in talk of the internal perspective, especially when considering causality, it seems that the following three assumptions need to be added:[30]

[SUBSUMPTION] SUBSUMED BEINGS

We as beings are fully subsumed under some of the states;

[RELIABLE EA] RELIABLE EPISTEMIC ACCESS

We have reliable epistemic access to those states in **[SUBSUMPTION]**;

[M-COHERENCE] MENTAL COHERENCE

Our mental content forms a coherent narrative through time.

Subsumption here is taken to mean that all aspects of our being - our bodies, phenomenal mental states and such like - occur within and are fully part of the universe and so are, in essence, ontic aspects. How they are described and talked about is a separate matter. And we are assuming here that 'reliable' means there is some consistent relationship between the universe and the epistemic access we have.

These three axioms and three assumptions must, according to the Archimedean constraint, hang together nicely. There must be no potential tension between them and the assumptions must sit comfortably with the axioms or necessitate the addition of further, non-contradictory, axioms.

As a contrast to what we shall argue, we first put the case that there is some tension between them: we shall put the argument that a Humean universe cannot support regularity naturally. That is, we shall first explain why **[M-COHERENCE]** and **[O-SEQ]** and **[NO-CONX]** are usually considered to be in tension; and then we shall examine why the argument that we have put - that Humean universe cannot support regularity naturally - fails.

[29]We use the term 'objects' loosely here.

[30]We call these 'assumptions' as we want to leave their status open to discussion.

The argument says simply that if we assume **[RELIABLE EA]** and **[SUBSUMPTION]** then **[M-COHERENCE]** must require that the coherent narrative of our mental content is reflective of a coherent narrative across time. And as there is no connexion between one moment and the next in Humean universes, this coherent narrative can only arise by serendipity and this is very unlikely. So that it is hard to justify holding **[M-COHERENCE]** with **[O-SEQ]** and **[NO-CONX]** without placing a great additional burden on serendipity.

We might try to mollify this tension by requiring that mental states do require regularity through time in some manner. We could do this by adding that mental states can only arise when there is regularity across [long] sequences of temporally juxtaposed states. So that we could only possibly be present in a universe that did indeed exhibit regularity, regardless of whether that regularity was the result of mere serendipity or something more. This is really an anthropic style argument similar to some of those found in Barrow et al. (1988). We shall say a little more on this below.

However, the tension is illusory and we now put our argument as to why: **[M-COHERENCE]** is entirely consistent with **[O-SEQ]** and **[NO-CONX]**. The reason for thinking that it is not stems from a misreading of **[M-COHERENCE]**. In reading **[M-COHERENCE]** we have assumed that coherence of the mental [epistemic] states through time must be consistent with the time ordering of the states in the Humean universe. We have assumed that our *experience* of time, which is an internal perspective, must be the same as the Archimedean ordering of time. We have mistaken one perspective for the other and become muddled. We have smuggled into our arguments an axiom that is, in fact, not there and then used it to ill effect. If we require only that each moment in the time we experience appears somewhere in the Humean universe then that is adequate to allow **[M-COHERENCE]** to sit comfortably with **[O-SEQ]** to **[ISOLATION]**. This is because, for each moment where the mental exists, the mental must contain within it a memory of past deeds and an anticipation of an instantaneous future - one that is 'similar' to the present and consistent with the past.[31] It must contain these if it is to be true to **[M-COHERENCE]**. If each moment obeys this then the external order in which moments actually occur will not be relevant. They will have a second [intrinsic] ordering set by the mental content at each moment. So long as the mental content present at each moment appears *at some time* within the sequence of moments in a Humean universe then we, as beings only having an internal perspective, will perceive a continuity to time across these moments, whether these moments are actually ordered the same way or not in the external actual time.

What makes this possibility gain traction is that there is, by axiom, no connexion or influence from one moment to any other in a Humean universe. This means that there is nothing inherited

[31] Recalling our discussion on duration and dynamic states in section 2.4 above.

from previous moments and nothing that passes down to future moments. An inhabitant of any given moment can have no epistemic access to what is to come or what has been. As such, any claim that the ordering of temporal experiences that we have as inhabitants of the universe is the same as the Archimedean ordering of the moments must, if held to be true, be accomplished by an explicit extra axiom in our model. That is, **[M-COHERENCE]** can be read as having time as an internal feature and so sit comfortably with the other requirements, or it can be read as the claim that experiential time and Archimedean time are identical, in which case it stands in some tension. We can take **[M-COHERENCE]** as saying either:

> **[EPI M-COHERENCE] EPISTEMIC MENTAL COHERENCE**
> Our mental content forms a coherent narrative through EPISTEMIC [internal] time;

or as saying:

> **[ARCH M-COHERENCE] ARCHIMEDEAN MENTAL COHERENCE**
> Our mental content forms a coherent narrative through ARCHIMEDEAN [external] time.

We prefer **[EPI M-COHERENCE]** over **[ARCH M-COHERENCE]**, particularly as we can have no purchase from our internal viewpoint as to which time we are experiencing. As such, we may accept **[EPI M-COHERENCE]** and dismiss worries that arise from a need to explain the tension between serendipity and coherence.[32]

We can get an intuitive feel for this idea of intrinsic ordering by considering a pack of playing cards with some extra structure:

Imagine a set of twenty-six ordinary playing cards, perhaps hearts and spades, additionally each labelled randomly by different letters of the alphabet and the numbers one to twenty-six. A card picked at random may be, say, the seven of spades and also carry the number 11 and the letter P. If we shuffle these cards and place them in a pile then the cards have an order given by their order in the pile. The top card can be thought to be followed by the next card down in the pile and so on. This ordering is derived directly from the actual positions of the cards in an external space. However, the cards also possess intrinsic orderings which are not, most likely, equivalent to this pile-ordering. For instance, the cards have an ordering they derive from the numbers one to twenty-six, and another from the letters of the alphabet, A to Z. These intrinsic orderings are not realised, in most cases, by the pile-ordering but that does not detract from their presence. If we take this idea across to the states of a Humean universe we can see they that have an ordering derived from their presence in the external time in which they sit - somewhat like the spatial organisation of the

[32] It is interesting to compare intrinsic time with the ideas suggested by Barbour (1999) as a possible interpretive framework for quantum theory.

cards in our pile. Additionally, however they also possess natural intrinsic orderings that derive from the presence of mental states. For instance, this might be achieved by assuming every mental state has an immediate memory of both its immediate past and an anticipation of its immediate future, both combining to define its place in the sequence of moments derived from these mental aspects.

The possibility of **[EPI M-COHERENCE]** is perhaps surprising: it means that any given universe that we have epistemic access to is compatible with a vast number of Humean universes. All that is required is that each state that contains [a being with] mental content appear at some place in the sequence. That is to say, if we take it that mental content must either itself possess or be accompanied by some form of regularity from moment to moment, then there is no need for this ordering to be reflected in the actual ordering of the states in the Humean universe.

Accepting **[EPI M-COHERENCE]** suggests that perceived regularity is inevitable in a Humean universe that subsumes mental content if we take mental content to be marked by a consistent narrative in time. And further there are a myriad of possible Humean universes that are compatible with any given epistemic ordering of states. This rather takes the wind out of arguments that claim a Humean universe stands in tension with the regularity we perceive around us. For instance, on this assumption, we must reject Strawson's claim:

> In fact, experience of regularity of succession is evidence for the existence of objective forces in rather the same way as experience as a whole is evidence for the existence of an external world of objects. (Strawson, 1987, p. 264)

We now return to look at **[RELIABLE EA]**. We shall argue that it is this assumption, rather than regularity, that causes a tension between the axioms and assumptions for a Humean universe. Recall, **[RELIABLE EA]** says, in simple terms, that what occurs in our minds reflects, in some manner, the world around us. And, of course, if we are to maintain that the Humean universe is just about local matters of fact, we must ask how the information about this purported external world is 'carried' to our minds - what justifies the claim that our perceptions and concepts reflect something about the world around us. There is no deduction from the first three axioms to claims that what we perceive of the world should reflect the external world: there is no connexion between states, by **[NO-CONX]**, and no relationship between parts of each moment by **[ISOLATION]**. This is a fundamental flaw in the Humean set up.

It is this, **[RELIABLE EA]**, that provides the challenge to Humean universes. The challenge does not come from **[M-COHERENCE]**. We then see an advantage of adding a connexion to our metaphysics: it gives us some purchase on arguing against the need to add **[RELIABLE EA]** as a metaphysical axiom.

Further, the arguments we have put here suggest that the perceived regularity of the universe cannot be a fatal blow to the possibility of a Humean universe - in fact the contrary is the case as we can accommodate this perceived regularity by adopting **[EPI M-COHERENCE]** quite consistently. The difficulty arises for the Humean in claiming **[SUBSUMPTION]** and holding on to **[RELIABLE EA]** . And in the Humean case we could resolve this tension by adding **[RELIABLE EA]** as an axiom. But, perhaps a better way to resolve the difficulty of **[RELIABLE EA]** is to admit a connexion between states.

2.7 Completing the axioms for a connexion-bound universe: the Single-Time Axiom

Above we argued that **[RELIABLE EA]** must be taken as an axiom for Humean universes but can be derived in a connexion-bound one. Now we turn to argue that we should add an extra axiom to our model of connexion-bound universes and explain the consequences of adding it.

We have not forbidden intrinsic ordering in connexion-bound universes and it remains the best explanation for regularity regardless of universe type. Recall, that we need have no link between regularity and determinism; that is, ontic necessity does not have to lead to regularity. However, it does not seem right to allow intrinsic ordering to remain if the impingement of the universe and previous mental states on our current state is mediated by a connexion. Here we defer to intuition and claim that it is natural that, in a connexion-bound universe, epistemic time coincides with Archimedean time. We are thus claiming that for a connexion-bound universe the order in which we experience our thoughts and the machinations of the world impinging upon us is reflective of the order in which they actually happen in the universe. Our claim is one of intuition and so is, of course, not water-tight: there may well be arguments that go the other way. It may be possible to claim that the universe conspires to make temporally juxtaposed states somewhat different and that they still have, in consequence, an intrinsic order that is separate from their Archimedean order. To counter any such argument without engaging with them we choose to add our claim as an axiom rather than merely hold it as being an intuitive consequence of other axioms. We require:

> **[STA] Single-Time Axiom**: In a connexion-bound universe epistemic time must coincide with Archimedean time.

However, adding this axiom furnishes us with a new difficulty. We have already asserted that **[RELIABLE EA]** is best dealt with by introducing a connexion but we must now explain how we might unpack **[M-COHERENCE]** upon adding **[STA]** to the mix. The problem is that, whilst previously the coherence in **[M-COHERENCE]** could derive from intrinsic ordering and so placed little constraint on the order of individual states, now **[STA]** does place a constraint on the order of mental states: mental states must be ordered by the universe directly. This means we must, if we take it that the mental states exhibit a regularity through time, seek to explain why this regularity obtains. It is easy to require that the nature of the connexion forces regularity onto the universe and hold it is this very regularity that is then reflected in **[M-COHERENCE]**. That is, we might seek to deduce **[M-COHERENCE]** by positing that connexions must constrain the way the world evolves. However, we have placed no such constraint on the connexion. Even a deterministic universe need have no regularity. Ontic necessity does not carry with it any assertion of regularity. Rather ontic necessity merely makes an assertion about the type of relationship between moments, not the details of how it unpacks.

What does this mean for our model? We find that epistemic regularity now stands in some tension with the structure of our connexion-bound universes if we add **[STA]** but remains consistent and unproblematic for Humean ones under intrinsic ordering. Epistemic Regularity remains plausible for Humean universes as we may invoke **[EPI M-COHERENCE]** but, as the order of the states is battened-down by **[STA]** in a connexion-bound universe, no such recourse to **[EPI M-COHERENCE]** can be sanctioned. So regularity is actually a problem for connexion-bound models and not for Humean ones - the very reverse of what is usually taken to be the case, as witnessed in the Strawson quote above - recall, we quoted:

> In fact, experience of regularity of succession is evidence for the existence of objective forces in rather the same way as experience as a whole is evidence for the existence of an external world of objects. (Strawson, 1987, p. 264)

If we are to be content with holding to both the coherence of the sequence of mental states and to our connexion-bound model with **[STA]** then we must seek explanation or admit defeat by deferring to serendipity. There are, perhaps, two ways we can attempt to avoid deferring. One is, as we suggested above, to introduce a constraint on the universe that derives from the nature of the connexion itself. We might require that connexion-bound deterministic universes must always be regular. The other is that we can, perhaps, invoke anthropic arguments.[33] Anthropic arguments say, roughly, that we can only persist as sentient beings in a connexion-bound universe that is in some sense regular and so should not be surprised to find ourselves in such a universe. Here we shall not adopt the first option, of allowing the regularity to be a facet of connexions, as we see no *prima facie* reason that connexions should always behave in such a way in deterministic settings.

[33]See Barrow et al. (1988).

Instead, as we are to assume - at least for this study - both that we live in a connexion-bound universe and that **[STA]** holds then we choose to nod towards the anthropic option, albeit slightly tentatively. We emphasize that the move is not one that can be made in general and we do invoke an anthropic principle here only because it has explanatory force.

2.8 Epistemic access

We have argued that reliable epistemic access, **[RELIABLE EA]**, is a deduction from other axioms for connexion-bound universes and an axiom in the Humean case. We have also argued that regularity alone is more likely to derive from a Humean than a connexion-bound universe. In this section we consider whether our epistemic access is sufficient to discriminate between universe types.

In this study, we shall hold that we inhabit a connexion-bound universe; or at least we shall be investigating the possibility of a sterile form of free will within such a universe - as it gives us scope to talk of both determinism and indeterminism. However, whilst the arguments and definitions above clarify the assumptions we are making for our models, it is useful to ask if there is anything in the models that suggests we might be able to tell what sort of universe we actually inhabit. Here we answer the question by introducing the claim as an additional axiom and then discussing some of its implications.[34]

Recall that both the Humean model and the connexion model both have a time ordered set of states but they differ in that the connexion model has additionally a connexion between the states. In our case we have the connexion as an ontic feature of the universe itself. However, we make it an assumption of both models that we only have access to the states themselves. Formally:

> **[EA] The Epistemic Assumption**: We only have possible epistemic access to the content of universe states.

That is to say we have no manner of distinguishing between a Humean universe and one with a connexion as all our knowledge derives from states alone. We also take it that mental states exist solely subsumed within these universe states but are mediated temporally by the connexion.[35]

It is useful to see where this takes us and what it permits in argument. Under this assumption we can take any observation we make and consider it both from a Humean and a connexion-based perspective. We can even go further if we accept the idea of intrinsic ordering and then pair each given observed universe, considered as a set of observed states, with a large set of Humean

[34] We return to this in section 5.13 when we revisit **[EA]** in the light of immanent relata.

[35] See also our discussion following Theorem 1 in section 5.5 and our discussion in section 6.4.

universes. For instance, consider the set, B, of *observed* moments between two times:

$$B = \{\mathfrak{U}_{obs}(t_\alpha) : t_m \leq t_\alpha \leq t_n\} \tag{2.8}$$

Where the subscript on $\mathfrak{U}_{obs}(t_\alpha)$ emphasises these are *observed* states. Then there is a set, $\mathfrak{H}(B)$, of possible Humean universes in which this sequence could occur as an intrinsically ordered set:

$$\mathfrak{H}(B) = \{\mathfrak{H}_i; i \in Index\} \tag{2.9}$$

defined via:

$$\forall \mathfrak{H}_i \in \mathfrak{H}(B), B \subseteq \mathfrak{H}_i \tag{2.10}$$

Equally, the set B may be within a deterministic or indeterministic connexion-bound universe. **[EA]** tell us there are no observations we can make to decide which type of universe we inhabit. Indeed, in discussions we may well be able to pair epistemically identical universes with different connexions together, as there is no guarantee that identical moments are related by identical connexions. This is something we revisit below, in section 2.10, where we discuss our notion of 'the laws of nature'.

As an example, I may watch a billiard ball move across a table. In a deterministic universe my observations are formed from a concatenation of mental states that are joined to the external event of 'billiard ball moving along' by a connexion. Under the epistemic assumption there might be a Humean universe that mimics this universe both in terms of the position of the billiard ball as it moves across the table and in terms of my corresponding concatenation of mental states. In the Humean case we have no connexion joining the mental states together, we just have either serendipity or some intrinsic ordering, and the relationship between the billiard ball's moving and my mental states is left as an assumption, expressed axiomatically as **[RELIABLE EA]**. The epistemic assumption means that there is nothing in the experience of observing a billiard ball moving across a table that permits us to distinguish between our being inhabitants of a Humean universe and our being inhabitants of a connexion-bound universe.

2.9 Predictability

In this section we look at how the notion of predictability fits into our models. We do this because we do not want to confuse the issue of predictability with that of determinism in our discussion of free will. We explain here why they are distinct and inequivalent notions.

Predictability is often confused with determinism.[36] In this section we discuss predictability from both the Archimedean and internal perspectives and in relation to both Humean and connexion-bound universes.[37] This section serves to emphasize the differences between ontic determinism and predictability and so places restrictions on the way we introduce ideas of predictability into our discussion of free will.

We start with two definitions of predictability:

DEFINITION 2.8: Archimedean Predictability

A universe is **Archimedean predictable** if and only if it is ontically deterministic.

DEFINITION 2.9: Internal Predictability

A universe is **internally predictable** if its future states are fully amenable to complete description using a fully articulable theory and at most a complete description from current states.[38]

We can think, somewhat informally, of the first definition as saying that a universe is Archimedean-predictable if we could gain a position outside of it and we could access both the complete details of any state and then, given we also possess full 'knowledge' of the nature of the connexion or access to some inventory that lists its features, we could construct or calculate the details of the next state of the universe.[39] Whilst this sounds like an epistemic construal of determinism, it is not. The definition itself characterizes Archimedean predictability in full and our characterization of it in Laplacian like terms is merely an explanation of the motivation behind the definition. It is important to note here that Archimedean predictability does not require any patterns or regularity in the universe: there is no restriction that the relationships set up via the connexion be the same, similar or related from one moment to the next. As such, when we talk of 'knowledge' of the connexion we must realise this is trash-talk - we mean by it only that a being outside of the universe could, with a complete inventory, construct future states and the construction may be no more than listing or it may be akin to our conception of calculation using mathematical schemata. It is also important to note that a Humean universe taken as a time ordered sequence of states is not Archimedean predictable. The question as to whether the epistemic ordering that arises from

[36] The assertion of confusion is valid only when unpacking determinism as an ontic doctrine and not an epistemic one. Of course, if we choose to associate determinism closely with [scientific] theory then there is no confusion.

[37] We do not look at probabilistic ideas of predictability here. This is partly because we hold they must involve immanence - see Chapter 4 - if they are to be reflections of more than serendipitous statistical regularity; and also because probabilistic notions play no explicit role in our arguments.

[38] We could modify this definition to include probabilistic predictability by replacing 'complete description' by 'probabilistic description' and then fleshing out the nature of what probabilistic prediction is to mean. We choose not to do so here.

[39] 'Complete' can be replaced by 'some sufficient complete subset' if we allow for some parts of the universe to have no causal consequents.

the intrinsic nature of inhering mental states could be considered to be Archimedean predictable is left as an open question: we leave the question as to whether we can remove the ontic aspect of determinism but still have some force of necessity between states in such a sequence of states - a necessity deriving from the intrinsic ordering. We have inserted the ontic restriction into the definition to exclude this possibility for the remainder of this work.

Internal predictability is more complicated. In simple terms it involves our epistemic access to states only and makes no demands on determinism or connexions. To be internally predictable requires the universe to be very regular and it is primarily an epistemic notion. We need to be able to identify categories of aspects within the universe at each moment and across moment. These categories must form, when complete at any time, a sufficient description of the universe to determine, through mathematical calculation, the exact and complete details of the next state in the sequence. This is then verified by the actual occurrence of a state that concurs with the prediction. We require the theory to be 'articulable' and by that we mean to say that it is accessible by us or by some suitable calculating device.[40] And we use the term 'complete' to indicate that there are no probabilistic notions involved, so as to rule out, for instance, the sorts of probabilistic calculations from some quantum theory schemata.[41] This definition makes one, impossible, request on any inhabitant of the universe. The impossible request, which is not a serious breach here, is that any being who persists within a universe can then have full access to the state of the universe at any time. In practice, we do not seek to predict the evolution of the whole universe, just small isolated parts of it under idealised conditions.[42]

The first definition of predictability does not require the universe to be regular as there is no restriction on the nature of the connexion or the states from moment to moment - determinism places restrictions only on the nature of the relationship between them, not on their internal constitution. The second definition of predictability equally could arise in a universe that is either Humean or connexion-bound and non-deterministic, in the ontic sense we claim here. That is, the assumed synonymy between determinism as an ontic concept and predictability is a false one.

To draw a contrast between predictability and determinism a little more starkly, we can offer an alternative definition of determinism, based on epistemic grounds, as follows:

[40] We leave such a calculating devices a little vague here. We might think of it as a universal Turing machine.

[41] We could extend the definitions to probabilistic predictions, such as those often furnished by quantum theory. We choose not to as it is not central or relevant to our arguments. Further, probabilistic predictability can also be taken across to the Archimedean realm using immanence as an underpinning. We introduce immanence in Chapter 4 but we shall not pursue probabilistic interpretations of Archimedean predictability there or here.

[42] We put aside issues of a 'Mary's view' here. See Jackson (1986).

DEFINITION 2.10: Epistemically Deterministic Universe

A universe is **epistemically deterministic** if and only if, every [suitable] finite subset of it is internally predictable in principle.

This definition is, admittedly, a little imprecise and would need polishing in any detailed study but it expresses the spirit of a common notion of determinism - those that take the idea of Laplacian predictability, ontically construed, across to internal predictability without comment.
What this definition is meant to say is that we can count a universe as being [epistemically] deterministic if it seems again and again, for all time, to concur with the calculations we make in predicting the future of [small] bits of it. And this does seem to be the way that most definitions of determinism tend to be formed or, at least, interpreted. This probably stems from confusion between the Archimedean and internal perspectives and a failure to apply the Archimedean constraint carefully. This is a purely epistemic notion of predictability and bears no direct relationship to the ontic distinction we make here between types of universe or the idea of ontic necessity.

2.10 The laws of nature

In this section we discuss what we are to mean by 'laws of nature'. We do this as we shall refer to laws of nature in subsequent chapters; there we shall mean ontic laws as opposed to epistemic laws and this section sets out the key differences between them in the context of earlier discussions.
First we unpack what we mean by epistemic laws and argue that we cannot make deductive moves from epistemic claims to ontic claims about connexions. We then move to unpack ontic laws as a contrast to epistemic ones. We then turn to argue for one feature that we shall require of ontic laws here: we require them to be 'essential' to the universe. In the third section below we shall start to outline what it is for ontic laws to be essential but we conclude the discussion at the end of Chapter 3. We postpone the completion of the discussion of essential laws until then as we need to marshal the technical devices we develop there to explain and analyse the notion in detail.

Before we begin we reiterate one point about probability: we shall not concern ourselves explicitly here or elsewhere with probability or probabilistic laws - although everything we claim concerning epistemic laws and epistemic regularity can be modified to accommodate such laws. However, understanding how to apply probability to ontic laws, and understanding what the probabilities might then mean in ontic contexts, is much more difficult. This is because, in part, ontic laws are our name for the machinations of the connexion in a deterministic universe. And in such a world there is no room for probability. If we do attempt to make room for probability within the ontic setting that interests us here - one where the problems of free will are glinting in the distance - then we encounter a difficulty: the difficulty is to understand what ontic construals of probabilistic claims about the future might mean in indeterministic connexion-bound universes. Such probabilistic claims are difficult to unpack if they are to refer to single moments in a universe's history. This is because it would seem, in accordance with the definitions and models we have outlined so far, the future has no ontic presence in the present in indeterministic connexion-bound settings. And if there is no future presence, it is difficult to unpack what ontic aspects the probabilities refer to in a manner that is both perspicuous and useful for what we intend here - namely an ontic construal of free will. As a detailed exposition of probability and its interpretation in the ontic realm is not directly relevant to our work, we make the decision to eschew any further explicit formal discussion of probability.

2.10.1 The laws of nature epistemically construed

We take epistemic laws to be systematizations of our observations. As such we broadly agree with Ramsey's encapsulation: Ramsey (1990, p 150) says that laws of nature are '…consequences of those [general] propositions which we should take as axioms if we knew everything an organised it as simply as possible in a deductive system.'[43] We deliberately take our construal of epistemic laws in the broadest possible terms: the important point here is merely that epistemic laws must take their cue solely from our observations initially. Whether these be direct observations or vicarious ones, made through devices of some sort, is not relevant. How we systematize our laws is not relevant nor is the issue as to how we relate the schema in which we express the laws with the world that we observe.[44]

We now examine a case for one conclusion: that we have no grounds within our model to make the move from epistemic laws, no matter how generously construed, to ontic claims concerning connexions. This is important for what is to be argued in future chapters - especially Chapter 4 - and it fits closely with our earlier claims concerning epistemic regularity and determinism. In Chapter 4 we shall investigate what it means for a universe to present 'choices' and there the schism between the ontic and the epistemic will play a vital role in our march towards our introduction of immanence into indeterministic connexion-bound universes. Immanence will play a crucial role in our proposals for an ontic unpacking of free will. Here we advance the case for our claim that there is no deductive move in two distinct stages: first we argue formally, using the machinery we have introduced earlier, explaining why the move from epistemic laws to ontic

[43] We do not engage with what 'organised as simply as possible' might mean.

[44] There is vast and sophisticated literature on laws of nature which we then bypass here. Some useful studies include: Suppe (1977), Dretske (1977), Cartwright (1983), Armstrong (1985), Swartz (1985), van Fraassen (1989), Bigelow et al. (1992), Loewer (1996), Lewis (2001a), Bird (2007), Ellis (2007) and Tegmark (2008).

claims is invalid; and second, we give a specific example drawn from the plethora of interpretations of quantum theory to illustrate some related points involving determinism, indeterminism and probabilistic theories.

Recall, above we claimed:

> **[EA] The Epistemic Assumption**: We only have possible epistemic access to the content of universe states.

We shall construct the strongest possible epistemic laws couched as a mathematical theory and explain why it cannot be used to make ontic claims about connexions across time and hence cannot be used to make claims as to ontic determinism and indeterminism. Some of what we discuss holds the seeds of the approach to causality we shall take up in Chapter 3 and the argument can be read again after that chapter. We shall assume we inhabit a regular connexion-bound universe and set aside the issue as to whether it is deterministic or not. We shall also assume that our probing of the world has allowed us to identify the fundamental constituents of the world perfectly - they would be our fundamental relata of Chapter 3 - and note that this is in keeping with **[EA]**. And further, admittedly stretching credence a little, we assume we know with certainty that these are indeed the correct fundamental constituents, leaving aside issues as to how we might have obtained this miraculous knowledge.[45] Let us assume further that we have identified patterns in these fundamental constituents as the universe evolves. So it might be that we have noticed fundamental 'object' A seems to interact with fundamental object B and when it does it always produces fundamental object C. We shall assume that the patterns we have noticed persist throughout all time and can be articulated by some finite mathematical theory that we also have worked out in full detail.[46] We have, in essence, a perfect theory. It relates without fault the patterns that we observe from moment to moment: it predicts with a perfect record of success. The question we ask is, then: why can we not make any ontic claims about connexions from this perfect epistemic theory?

To explain the answer we must first outline what we are to mean by ontic claims about connexions. Ontic claims involve two aspects: we must be able to claim that certain fundamental objects at one time are related to fundamental objects at another via a local connexion; and second, we must be able to make claims as to whether each connexion is a necessary one or not. We can make neither claim. First, we cannot make claims that the mathematical [or other] relationships we have worked out between the constituents of our world, the 'objects', reflect the actual relationships between them. There may be other mathematical schemata that give the same predictions but pose

[45] We make the second assumption because our arguments do not concern this knowledge but rather concern what we can say about trans-temporal connexion relationships.

[46] Again we put aside difficult issues such as those that derive from our being within the universe and the calculations' occurring within the universe too.

different relations between the universe's constituents. Or there may be no finite theory at all that summarises the actual connexions - it might just be pure serendipity that the actual machinations of the connexions across time, even if it they are necessary ones, produce the results we observe. And additionally we have no way of telling from our observations whether the relationships from moment to moment are necessary or a mixture of necessity and serendipity or the result of pure serendipity. In simple terms we might be observing the sequence of positions in a game of chess where each move follows the normal rules as far as we are concerned but all whilst the actual players are using different rules or playing with reckless abandon.

We add one further point here: our argument is strengthened to the point of obviousness if we assume that we inhabit a Humean universe. In such a universe there is no move to be made from perfect predictability to assertions about connexions. And further, recall we argued above that we cannot tell whether we inhabit such universe or not.

Let us now turn to an example from quantum theory that illustrates some additional points.[47] The example shows that mathematical theories alone are inadequate as we must link them to the world and they must carry with them some interpretive baggage, governing the way we link theory and world, which may lead in differing directions.

The original interpretation of the mathematical schema of quantum theory, both Schrödinger's wave mechanics and Heisenberg's matrix mechanics, suggested that the world is inherently unpredictable.[48] The 'inherently' is important here - they interpreted their mathematical schemata as modelling the universe as it really is and so took it that the current state of the universe does not uniquely determine future states. This made those wedded to a certain sort of determinism uncomfortable and there were suggestions that the theory as it stood was incomplete - in that there was more mathematics to come that would reflect a deterministic world and that this extra mathematics represented extra ontic features of the world hidden in some way. But von Neumann proved that the addition of extra hidden structure in the mathematics - and hence by extension ontic features of the world - to quantum theory would not resolve the problem and his proof was taken as the final settlement of the matter for some time.[49] However, an alternative way of dealing with the mathematical scheme, originally proposed in one form by de Broglie,[50] was resurrected by

[47]A non-technical account of the historical background and development of quantum theory can be found in Kumar (2009). A more technical examination of the hidden variable approach can be found in Bohm & Hiley (1993) and some context in Jammer (1974, Ch 7).

[48]These two mechanics are, in a certain sense, equivalent mathematical schemes. The original proofs of this are found in Eckart (1926) and in Schrödinger (1978).

[49]The proof can be found in von Neumann (1996).

[50]The idea was proposed in his thesis and appears in de Broglie (1925).

Bohm.[51] Bohm demonstrated that von Neumann's proof was incorrect as it made an assumption. It assumed the hidden variables needed to complete the quantum theory represented, in simplistic terms, spatially-local matters. Bohm showed that it was possible to rewrite the mathematics of the theory in such a way that it became deterministic - as a theory - but his reworking was at the expense of making it into a [non-local] hidden variable theory.[52] This non-locality was something that von Neumann's proof did not anticipate as he had implicitly assumed that hidden variables, as ontically construed, acted 'locally'. The result is that we can have a quantum theory that deals in probabilities and adopts the assumption that unpredictability in a theory reflects indeterminism in the universe - something we have shown to be in error as a deductive move. Or we can have a theory that eschews locality in a certain sense but does not postulate inherent probabilities - and so, again on the mistaken assumption that epistemic predictability is equivalent to determinism, conclude that the universe is deterministic.

The point is that the mathematical schemata and the interpretative baggage mix together to produce alternative possibilities. It may be that out universe is amenable both to those theories which give full predictably but eschew such sacred doctrines as locality, temporal ordering and so forth and also to theories which give probabilistic predictions at best but cling to shibboleths such as locality.

2.10.2 The laws of nature ontically construed

Let us now turn to the second use of 'laws of nature', namely the ontic construal of the term. Here we link it with deterministic connexion-bound universes. We take the term 'laws of nature' to be shorthand for talking about the part the connexion plays in the binding of states one after the other in deterministic universes.

There are some points we mentioned earlier that are useful to recall here: first, we took it that the connexion provided the necessity that made one state inevitable on the back of another in deterministic universes - so deterministic universes, by definition, are connexion-bound and so are subject to ontic laws of nature; second, the presence of a connexion does not lead to determinism as it may not make one state necessary on the back of another - so the mere presence of a connexion

[51] See Bohm (1952a), Bohm (1952b) and Bohm & Hiley (1993).

[52] It is not clear if Bohm originally intended his non-local interpretation to represent something ontic as well or whether he was merely exploring alternatives to the orthodoxy. He was certainly looking to produce a deterministic view of quantum theory and was encouraged in his endeavours by private meetings with Einstein. Certainly in later work, Bohm and his co-workers took the view that the quantum potential, essentially the mathematical referent of the hidden variable, had ontic aspects - even using the term 'Ontological Interpretation' for his ideas in Bohm & Hiley (1993). In passing, it is worth noting that many discussion of the quantum potential suggest that it is an addition to the theory: it is, in fact, already present in the standard mathematics of the original version of quantum theory and all that is required to make this explicit is a rewriting of the underlying governing equations in a slightly different, mathematically equivalent, manner.

is not sufficient to make claims that a universe is subject to ontic laws; third, we argued above that there is no move from determinism to regularity - that is, there is no reason to require that a connexion, even when giving us a deterministic universe, forces behaviour that is regular from place to place or from time to time. A deterministic universe may appear entirely random, or highly organised or as anything in between.

Throughout this study we shall be concerned primarily with ontic laws. We shall use the terms 'laws' or 'laws of nature' to mean 'ontic laws' henceforth. If we wish to make a contrast with epistemic laws of nature we shall be explicit and use terms such as 'epistemic laws' and 'ontic laws'.

2.10.3 A constraint on ontic laws

For the ontic laws, we shall require a structural constraint. We give an informal outline of the constraint here and continue our discussion in more detail in sections 3.9 and 3.10.

The structural constraint we make is to require all ontic laws - the way the universe is forced to evolve deterministically - be set in advance and held in the very fabric of the universe. We hold that these ontic laws are held in the very fabric of the universe in order to ensure ontic necessity arises from the ontic nature of a universe state itself. That is, when we have ontic necessity operating from one state to the next, expressed in (2.1) as $\mathfrak{U}(t_n) \boxdot\!\!\rightarrow \mathfrak{U}(t_{n+1})$, the necessary way the state $\mathfrak{U}(t_n)$ evolves into the state $\mathfrak{U}(t_{n+1})$ is fully governed by the [causal] nature of the state $\mathfrak{U}(t_n)$. Further we can then say this nature must be 'set in advance' as we require ontic necessity to be transitive, as expressed by equations (2.2) and (2.3). Recall these equations said:

$$\mathfrak{U}(t_a) \boxdot\!\!\rightarrow \mathfrak{U}(t_z)$$

can be considered as shorthand for the sequence:

$$\mathfrak{U}(t_a) \boxdot\!\!\rightarrow \ldots \mathfrak{U}(t_m) \ldots \boxdot\!\!\rightarrow \mathfrak{U}(t_z)$$

So the nature of each moment in a deterministic connexion-bound universe derives by necessity from the previous moment. It is in this sense that we mean that the laws are 'set in advance'.

Whilst we call this a constraint we argue that it must be the case in deterministic universes. Indeed we think of ontic laws as set at the beginning of the universe.[53] This does not mean the ontic laws need operate in the same manner moment to moment or place to place but it does mean the way

[53] We assume a beginning to simplify our discussions. Otherwise, we must deal with how our definitions fit with an infinitely extending past.

the universe transits from 12 pm to 12.01 pm on 13th January is determined in advance - that is, the universe must carry along in its very fabric the ontic laws. Later, in sections 3.9 and 3.10, we shall unpack this further and show how ontic determinism means the laws must be, in a certain sense, essential to the universe. If we do not require this of ontic laws and instead allow them to come into being at each moment without prior necessity then our distinction between determinism and indeterminism collapses for universes.

A digression to the Humean case may help to clarify this. Imagine a billiard ball moving along in space. At each moment we expect it to continue to move along unencumbered unless acted on by some external resultant force - that is whacked by something. Now imagine that whilst we are watching this billiard ball it suddenly makes a sharp right turn and then continues in a new straight line. We might say that it has violated the laws - which must of course be epistemic here. This is not really a correct assessment of the situation. Rather, as the laws in the Humean case cannot be set in advance, all we have is an example of a new phenomenon that must be added to the remit of the epistemic laws as they stand to produce a new, perhaps wider, set of epistemic laws. That is, in a Humean universe, laws must conform to whatever happens and the more things that happen, the more the laws must accommodate: that is, events are not subject to laws but laws are subject to events.

Now, consider a deterministic universe with a connexion. Recall that the connexion guides the universe from moment to moment; it binds states together. If we allow the nature of connexions to be set only at the moment of their operation - so they do not inhere within the current state $\mathfrak{U}(t_0)$ - then they cannot be predetermined by anything that has gone previously and can no longer make future states necessary on the back of previous states. The only other option, if we are to maintain determinism as ontic necessity, would then be to allow the nature of the evolution to be set *externally* to the fabric of the universe - something we examine in sections 3.9 and 3.10.

How far in advance is acceptable? Could we have them set a minute in advance consistently, or two minutes or must they be set from the outset - whatever the outset of a universe might be? The answer is they must be set from the outset and carried along in some manner with the universe. If they are not set in advance then it is not true to say that the state of the universe at any one time has a unique future: this is because any ontic laws set during the universe's evolution invalidate the notion of ontic necessity.

We return to this in sections 3.9 and 3.10. There we shall look in more detail at what we mean by our essentialism and what this claim is to mean.

2.11 Looking forward

In the next chapter we start to take apart the internal workings of our model universe. From here onwards we shall assume that we are placing our arguments within a connexion-bound universe and we shall assume that time is Archimedean. We shall make no claims as to whether our universe is deterministic or not - rather, we leave that as an open question. Humean universes now take a back seat.

To summarise what we must carry with us from this chapter:

We have two assumptions we shall make:

- We inhabit and are fully subsumed within a connexion-bound universe;
- Epistemic time is the same as Archimedean time.

And a number of notions that we have discussed:

- Determinism is an ontic doctrine;
- Deterministic universes ontically necessitate their own futures via connexions;
- Ontic necessity does not require there to be any regularity within a universe;
- Regularity and predictability are not a mark of ontic determinism;
- We explain the regularity of our observed universe by nodding to anthropic principles;
- We have only epistemic access to states not connexions;
- There is no move from epistemic laws to ontic claims concerning connexions;
- Indeterminism for connexion-bound universes is the antithesis of ontic determinism;
- The 'ontic laws of nature' is the name given to the machinations of the connexion in a deterministic connexion-bound universe;[54]
- Ontic laws are 'essential' to a universe - see sections 3.9 and 3.10;
- It is impossible for us to tell whether we inhabit a deterministic or indeterministic universe, or to differentiate between connexion-bound and Humean universes.

[54]Recall we shall use the terms 'laws' or 'laws of nature' to mean 'ontic laws of nature' henceforth.

Chapter 3

Causality

3.1 Introduction

In the previous chapter we outlined a view of determinism for connexion-bound universes. There we took a universe to be an ordered sequence of moments parametrized by a discrete set. However, we said very little about the micro-structure of the moments or about how the objects and 'things' that inhabit each moment relate, through the connexion, across moments. In this chapter it is our aim to unpack the micro-structure of the universe and to set out rigorously the relationships between moments. We achieve this by setting out a theory of causality. It will then be these theoretical tools that we shall use in our assault on free will, and supervenience, in the following chapters. And, for the moment, we take it that causality is the only ontic relationship between relata.

There are a number of approaches to causality in the literature: for instance, amongst many, counterfactual theories,[1] probabilistic theories, necessary connections and so forth.[2] Here we largely discount much of these alternatives without comment and instead build our own generalised model for causality based around the idea of a connexion at the level of relata. In a more complete treatment, we would aspire to extend our discussion to encompass a careful and extensive analysis of how our model fits with other theories and how it might be used to address and clarify difficulties found in these other models. We do demonstrate, albeit briefly, in an appendix to this chapter that our causal structures are related to causal sets - a mathematical scheme that has been proposed to deal with issues that arise in constructing theories of quantum gravity.[3]

We justify the lack of detailed comparisons with other models on a number of grounds:

First, we stated at the beginning of Chapter 1 that one aim of this study is to produce an ontic

[1] See Collins et al. (2004) for a survey.

[2] Surveys can be found in Beebee et al. (2012) or Paul & Hall (2013).

[3] A summary of causal set theory and its rationale can be found in Dowker (2006).

construal of the operation of free will. And so we must seek ontic construals of the metaphysical constructions we seek to marshal in achieving our aim. We must, then, seek an ontic model of causality if we are to use it as a route to achieving our aims. In Chapter 2 we argued that ontic determinism and epistemic predictability are not to be conflated. Theory-based predictability is an epistemic notion so any theory of causality that relies on predictability must run the risk of being one that reflects epistemic rather than ontic aspects of a universe. So we reject models that rely on solely epistemic criteria as plausible candidates for an ontic construal of causation.

Second, we hold as an assumption in this study that we inhabit a connexion-bound universe and not a Humean one. We set out some reasons why this is a preferable model to the contrasting Humean model in Chapter 2. As such it is natural to take the ontic connexions between states as being the ontic markers of how aspects of one moment lead to aspects of other moments. Further, we take causality to be a theory that tells us how one set of 'things' in the universe leads to the formation or presence of other 'things' at later times. As such the arguments and discussions of Chapter 2 strongly suggest that we should base a notion of causality around the connexions that bind moments with moments. And this then has the advantage of giving an ontic construal of cause and one that is not beholden to notions of regularity.

The idea that there is some connexion from moment to moment is not new and there are detailed philosophical models of causality and the like based broadly on such a notion.[4] We, however, make no attempt to compare the machinations of our model with others. Nor do we attempt in this chapter to address the criticisms that these other models have elicited. We do, however, review our model in Chapter 7.

We noted earlier that Hume can perhaps be read as denying a connexion - although as we also noted this is not clear - in contrast, Kant does seems to hint that he might embrace the general idea. Indeed he says,

> ...we can conceive a whole series of substances of which the first transmits its state together with its consciousness to the second, the second its own state with that of the preceding substance to the third, and this in turn the states of all the preceding substances together with its own consciousness and with their consciousnesses to another. (Kant, 1964, A 364, p. 342)

Kant is not talking strictly of causality in the form we have it here, as he is concerned with the propagation of the self [an 'I'] through time, but the general thrust of his idea here does resonate with our theory of causation.

[4]Notable discussions can be found in Fair (1979), Dowe (1992), Salmon (1994), Dowe (1995), Salmon (1997), Schaffer (2000), Dowe (2000). And some more recent discussions in Lupher (2009) and Bigaj (2012).

So, the framework for causality that we develop here takes a strongly ontic view of cause based on the local behaviour of connexions. We assume that causation is a real, objective feature of the world and offer a tentative general model for it. In particular, there is an assumption throughout that for one 'thing', or perhaps 'event', to cause another there must be some transfer or [ontic] influence 'flowing' between them. This might be a transfer of energy-momentum or a more abstract transfer of information but, whatever it might be, causality is at base an ontic connection [a connexion] across time. This assumption is not justified here but taken as being fundamental to any ontic outline of causality that is based around our models of connexion-bound universes.

Before we begin to set out our model, a few informal comments and reminders to set the background: first, we assume that the universe evolves in time - each moment is compelled to transform into or influence the next and we take this evolutionary process as basic. We do not seek to justify why the universe works this way or attempt to understand what time in this process is. We take it there is time and that the world marches forwards from moment to moment, creating each new moment out of the old.[5]

The second aspect we take as basic is that we inhabit a connexion-bound world and that there is an influence or transfer from moment to moment: moments are tied together in some manner by the connexion. We do not make an issue of whether each moment arises anew out of nothing, perhaps guided by the previous moment, nor do we make an issue of how the influence or transfer is to operate between moments and across time. We take it that it does and say no more.

Throughout, we hold that the mental - broadly construed - is fully subsumed within the causal schema we outline here. That is we assume that when we talk of relata and connexions then the mental is subsumed in this talk. We have already seen this in Chapter 2 where the relationships between our perception of the world and the world in itself came into play when discussing our universe models. We make a number of comments and examine a number of consequences of this assumption throughout this, and subsequent, chapters. However, we postpone a discussion of this assumption until Chapter 5. There, in section 5.5, we make the assumption explicit - as Theorem 1 - and explain both why we take it as an assumption and why we consider it to be a *plausible* assumption for our study.

[5] Later we make one exception to this when we discuss indeterminism in Chapter 4. There we countenance the possibility that, in an indeterministic universe, time might just cease in some manner. As this is a very minor part of our discussion in Chapter 4, and does not affect our arguments there, we see no harm leaving our assumption on time to stand here.

3.2 Causal frameworks and the causal pre-bundle

We now start to set out parts of our causal model. In this section we define and outline the structure of both a **causal framework** and of a **causal pre-bundle** of frameworks. It will be on the back of these structures that the causal relationship will be established and around them that our discussions will orbit.

At each time we have that the universe divides into a set of fundamental causal relata. These relata will then be the basis of the structure on which we define the causal relationship and start to build our picture of causality. As we seek generality here we shall set aside the issue as to what these causal relata are.[6] We assert only that they have ontic status and are independent of descriptions we may impose on the world. Of course, if we were to discover what these relata are, then we should expect them to be related in some manner to the world of experience. This relationship need not be as simple as is often expected. For instance it is not clear that relata need be expressible in terms of simple spatio-temporal regions. This thinking derives from issues that arise in interpretive aspects of quantum theory, such as an ontic [ontological] status of Heisenberg's uncertainty principle,[7] or that the formal mathematical structures of quantum theory are often constructed in phase-space and not space-time. Additionally, individual relata may not correspond to local regions in space-time in a trivial manner - for instance, both quantum entanglement and issues with dealing with energy in gravitational waves imply something more subtle is required.[8]

We start with the idea of a causal framework, or 'framework' for short. We take the universe at some moment in its evolution,[9] and hold that it divides uniquely into a finite or countably infinite set of fundamental causal relata. We explain why they are fundamental below. These are then used to generate all possible causal relata at time t. We then impose some restrictions on the way these relata combine under set operations. So, a universe at time t divides into a set of fundamental causal relata:

$$f\mathbb{C}(t) = \{f\mathcal{S}_i(t)\} \tag{3.1}$$

where we use the prefix 'f' to denote fundamental status and i is an index which we shall take to be discrete.[10] We shall use $\mathcal{S}$ throughout as our standard symbol for relata and, as mentioned, for fundamental relata when prefixed by f.

---

[6]For a discussion on relata, see Ehring (1987).

[7]See, for example, Jammer (1974) or Bohm (1989).

[8]For a brief discussion on 'Why the energy of the gravitational field cannot be localized' see Misner et al. (1973, pp. 466-468). Also Weyl (1999, §33, pp. 268-273).

[9]Here, recall, we take a naïve view of time.

[10]Here we note that aspects of modern physics such as Heisenberg uncertainty principle, thermodynamics, and the non-localizabilty of energy in general relativity suggest that we cannot reduce causal relata to points in some underlying space-time manifold.

Before we move to combining fundamental relata we shall say a little on sets and structured sets. A set here is to be taken as an unordered collection - usually of relata - whereas a structured set is a collection of elements arranged in some specific manner. Structured sets have members but also have some 'geometric' or structural aspects to them. In much of what follows we shall blur the lines between sets and structured sets. Sometimes we shall have structured sets of fundamental relata appearing in definitions and sometimes we shall just have sets of fundamental relata appearing - as mere collections. Which appears depends on the definitions in questions. We could attempt to be scrupulously notationally rigorous and define a map from structured sets to sets - mapping structured collections of fundamental relata to mere collections of fundamental relata - and then invoke this map in our definitions and equations at every relevant moment. In the interests of simplicity we choose not to do this. Instead, we shall talk of sets and structured sets without worrying too much about the distinction. What is meant should be obvious in each case and where there is doubt we shall write '[structured] set' rather than 'set' or add comments to our discussions. We justify this decision by noting that it makes much of what we have to say simpler to explain, without compromising clarity; and further we note that the structural aspects of the sets we define will play only some small part in the arguments we construct in future chapters.[11]

We allow combinations of fundamental relata, all taken at the same moment, into larger [structured] sets by means of union, $\cup$. We can then express the conditions for our set of relata to be fundamental:

DEFINITION 3.1: Fundamental relata

A set of relata is **fundamental** if it satisfies:

FR1: At each time the corresponding set of fundamental causal relata cover the universe, $\mathfrak{U}(t)$:

$$\bigcup_i f\mathcal{S}_i(t) = \mathfrak{U}(t) \tag{3.2}$$

where the union here includes a structural aspect.

[11] We avoid being specific about the structured nature of the set because it is not clear what sort of structural relations there may be between relata. For instance, it may be the case that sets of fundamental relata can co-occur within a single spatial region; or it may be the case that one fundamental relata has a presence at more than one spatial region. And further, it is an [implicit] assumption here that the spatial manifold itself arises *from* the relationships between relata, the structure as we have it, rather than the relata having a presence *on* such a manifold. We do not seek to justify this assumption here as it does not affect our arguments directly. Our assumption here that the spatial, or spatio-temporal, manifold is not primary is motivated by concerns from contemporary physics - see, for instance, Callender & Huggett (2001). Additionally there are some interesting discussions to be found in Seiberg (2005) and some discussion of discreteness and causal sets in Dowker (2006). With some simple additional definitions, our causal theory can be shown to be an example of a causal set theory - the basic axioms for these causal sets are outlined in Dowker (2006, p4). We show this, briefly, in an appendix to this chapter.

And also, allowing for intersections:

FR2: Any intersection of differing fundamental relata is empty:

$$\bigcap f\mathcal{S}_j(t) = \emptyset \;\; any\, j \tag{3.3}$$

Where **[FR1]** says the union of all the fundamental relata in the generating set when considered as a whole exhaust the universe.[12] Informally, this says that there is nothing in the universe at time t that does not occur within the fundamental causal relata or reduce to some [structured] combination of them.

And **[FR2]** says that fundamental relata are independent of each other in that they have no common causal content - something we say some more about below. They may be dependent in other ways such as having a common cause or being supervenient.[13]

We take it that a universe at each moment is completely specified by a unique set of fundamental relata structured, that is arranged, in some manner. So, given a moment in the universe's evolution - which we denote as $\mathfrak{U}(t)$ - we shall assume it is uniquely constituted of a set of fundamental relata structured in some manner. We call this a **causal framework** at time t. It should be noted that there need be no similarity between fundamental relata within or across causal frameworks and we remain neutral on this point.[14] And to reiterate: in this study, we shall make very little of the structured aspects of causal frameworks and most of our arguments assume the structure is there but make no explicit use of this.

For completeness, we define a causal framework with respect to a universe at some moment:

DEFINITION 3.2: Causal framework

A universe at any given time, $\mathfrak{U}(t)$ is constituted of a structured set of fundamental relata. Each given moment of a universe has only one set of fundamental relata associated with it and only one way of structuring this set. Such a structured set must obey FR1 and FR2 above and is called a **causal framework**.

[12] As we mentioned in footnote 11, there is more to this. Often it is thought that relata are expressible with respect to some manifold, such as space or space-time, but that this manifold has some independent existence. This is common, for instance, when dealing with mathematical models of physical phenomena - for instance phase-space is assumed as a manifold on which we can express our dynamical theories rather than taking phase-space as arising from the subject of the theories. Here we make no such assumption. Any manifold-type aspects of the universe must be held within the relata and causal structure in some manner.

[13] Supervenience is dealt with in Chapter 6. Note that Kim (2007) maintains that the supervenience relationship is both ontic and non-causal. Kim (2007, p. 34) says, 'supervenience ... is not a mere claim of covariation between mental and physical properties; it includes a claim of existential dependence of the mental on the physical.'

[14] In section 3.9 we say more on what constitutes identity and similarity of fundamental relata.

We can now form [non-fundamental] causal relata by taking various countable [structured] unions of fundamental relata. We define a general causal relatum at time t as a finite or countably infinite set of fundamental relata:[15]

$$\mathcal{S}_k(t) = \{f\mathcal{S}_j(t)\ some\ collection\ of\ j\} \tag{3.4}$$

All such combinations generate the set of causal relata at time t.[16] This leads to closure conditions on each causal framework. We assume the set of relata in any framework is closed under countable union and countable intersection:[17]

$$\mathcal{S}_i, \mathcal{S}_j, ... \in \mathfrak{U}(t) \Rightarrow \mathcal{S}_i \cap \mathcal{S}_j \in \mathfrak{U}(t) \dots \tag{3.5}$$

$$\mathcal{S}_i, \mathcal{S}_j, ... \in \mathfrak{U}(t) \Rightarrow \mathcal{S}_i \cup \mathcal{S}_j \in \mathfrak{U}(t) \dots \tag{3.6}$$

Considering the sets alone, we add one more condition:[18]

$$\mathcal{S}_i \in \mathfrak{U}(t) \Leftrightarrow \mathcal{S}'_i \in \mathfrak{U}(t) \tag{3.7}$$

If we consider these three conditions - (3.5), (3.6) and (3.7) - as conditions on the sets of fundamental relata, disregarding the structured aspects, then we can consider the fundamental relata in a causal framework as a generating set, under normal set operations. They generate, at each time, a σ-algebra at that time.[19] Of course, the exact nature of the σ-algebra can vary from moment to moment. This will be useful later when we come to place measures - quantifying maps - on our structures.

We can now define a **causal pre-bundle**.

> **DEFINITION 3.3: Causal pre-bundle**
>
> A **causal pre-bundle** is the totally ordered set of frameworks, ordered by their time parameters. Total ordering means we have a definitive idea of earlier and later. We have only to look at the time parameter value of two frameworks to tell which is earlier and which is later. $\mathfrak{U}(t_0)$ is earlier than $\mathfrak{U}(t_1)$ if and only if $t_0 < t_1$.

[15] We do not associate a general causal relatum with objects or events here. There are some restrictions that we shall introduced later as to what combinations constitute objects - for further discussion, see section 6.3.2.

[16] 'All' may be slightly generous - see section 6.3.2.

[17] Whilst these are listed as separate conditions they are interrelated by de Morgan's laws. With care, we can use countable union and complement to prove closure under countable intersection.

[18] The dash represents set complement.

[19] See Halmos (1974).

We have not, as yet, set up any relationship between the frameworks at different times in the causal pre-bundle. We cannot say that this pre-bundle alone constitutes a mimic of the universe from start to finish.[20] In essence the pre-bundle is just an ordered collection of [structured] sets, but sets without any defined relationships between them.

3.3 The causal relationship and the causal bundle

We now turn to define a relationship between these relata that merits the name 'causality'. We want to be able to talk of some relata as the cause of other relata in a manner that pays some heed to our intuitive notions. In this section we define the causal relationship and start to explain why it deserves its name. We also mention transitivity briefly - enough to clarify the latter sections of this study.

We start with the causal relationship between fundamental relata and then build causality on the back of this. We take it that a fundamental relatum at time t_0 can contribute, via a connexion, to more than one fundamental relatum at later times but that its contribution alone may not be sufficient to causally determine the later relatum in full. We discuss the idea of contribution below. We also take it that the fundamental relata are defined in such a way that the contributions of earlier relata cannot be broken down any further, so that fundamental relata do not have hidden causal parts. This in essence is to hold that fundamental relata are the smallest causal units. This avoids difficulties which may arise if we were to permit an infinite hierarchy of relata.[21] There is no concept of parts of fundamental relata at one time contributing to parts of fundamental relata at other times. We saw this idea earlier when we specified that fundamental relata do not intersect or have subsets that are causal relata. So given some fundamental relatum at time t_1, $f\mathcal{S}_i(t_1)$, we can define the cause of this relatum at some earlier time, t_0, as being the full set of fundamental causal relata that contribute to $f\mathcal{S}_i(t_1)$ - allowing that this set may be empty as in the case of spontaneously generated relata.

For the moment 'contribution' means that the causal set of fundamental relata at t_0 considered in isolation would evolve into - or influence, or cause via a connexion - the complete relatum $f\mathcal{S}_i(t_1)$ and perhaps an [incomplete] causal penumbra. We explain the penumbra below. We shall re-characterize the idea of contribution later when we introduce a measure structure onto frameworks.

The causal relationship sits on the back of the natural evolution of the universe from time t_0 to

[20] Again, for simplification we assume the universe has a well-defined beginning and a well-defined end. This assumption can be abandoned with almost no effect on the way causality is defined later.

[21] For a related discussion see Block (2003) and also Kim (2007, pp. 57-69).

time t_1 - and more generally from time t_n to time t_{n+1}. It is important to note that the slightly loose definition given here allows for both the possibility of causal over-determination and for the possibility of a degree of temporal non-locality. We shall tighten the structure later, when we define transitivity, so as to disallow temporal non-local causes.

To put the causal relationship more formally we define **causal maps** on our pre-bundle that give us the relationships between fundamental relata at different times. It is important to emphasize that these causal maps rest on the back of the connexion and natural evolution of the universe.
The causal map $\mathfrak{C}(t_1, t_0)$ maps $f\mathcal{S}_i(t_1)$, the **effect**, to a [usually un-structured] set of fundamental relata at time t_0, the cause of $\mathfrak{C}(t_1, t_0)$, or just **cause**, as per the discussion above.

$$\mathfrak{C}(t_1, t_0) : f\mathcal{S}_i(t_1) \longrightarrow \{f\mathcal{S}_j(t_0);\ \textit{some } j\,\} \tag{3.8}$$

Noting that if the spontaneous coming into being of relata is countenanced then, for relata with spontaneous relata in their history, we can only allow causal maps to extend back to the point when these relata come into being. We define how causal maps extend backwards in time in (3.12).
Rather than introduce further notation, we prefer to allow the right hand side of (3.8) to be thought of in a number of ways. We can think of it as either a structured set of relata or as just a set of relata - so ignoring the structured aspect. And further, we can think of the set as being made up from fundamental relata, or as one relatum constituted from the union of the elements of the set, or as being exhausted by some set of non-intersecting causal relata made up from various unions of the fundamental relata.[22] The exact choice depends on the context we are discussing.
As an example: if we have no causal over-determination then the union of the elements of the set will make up a single causal relatum which will then constitute the unique cause of $f\mathcal{S}_i(t_1)$. If we have the possibility of some form of causal over-determination then considering this set as a single relatum may be less useful so then we may opt to split it into a set of relata each of which is a sufficient cause. We shall define causal over-determination rigorously below.
We can define the reverse causal map of $\mathfrak{C}(t_1, t_0)$, written as $\mathfrak{C}(t_0, t_1)$, as the map that takes the cause of $f\mathcal{S}_i(t_1)$, namely $\{f\mathcal{S}_i(t_0);\ \textit{some } j\}$, to $f\mathcal{S}_i(t_1)$:

$$\mathfrak{C}(t_0, t_1) : \{f\mathcal{S}_i(t_0)\} \longrightarrow f\mathcal{S}_i(t_1) \tag{3.9}$$

Whilst we shall have little use for this, we note that it must obtain its interpretation from the original causal map. This is because we cannot talk of the evolution of the cause solely to its effect

[22]Later, in section 6.3.2, we do suggest there is some ontic restriction on which collections of fundamental relata can be considered as objects. The ideas we discuss there are not relevant to the comments here.

as we have the possibility of a causal penumbra.

Now, to explain what we mean by a 'causal penumbra': we see from (3.8) that the cause of the fundamental relatum $f\mathcal{S}_i(t_1)$ is given the set $\{f\mathcal{S}_j(t_0);\ \textit{some } j\}$. This set of relata will contribute to a number of fundamental relata at t_1 including, of course, to $f\mathcal{S}_i(t_1)$. Their contribution to $f\mathcal{S}_i(t_1)$ will be complete in that it forms [i.e. causes] the relatum in its entirety. The contributions that the set makes to parts of the universe outside of $f\mathcal{S}_i(t_1)$ at t_1 is what we term 'the causal penumbra of $\{f\mathcal{S}_j(t_0);\ \textit{some } j\}$ with respect to $f\mathcal{S}_i(t_1)$ at t_1'. We add 'with respect to' as we cannot guarantee that the causal set does not also cause other complete relata that constitute part of the penumbra here.

The causal maps then extend to general relata in a natural way by preserving combinations of fundamental relata by countable union and complement across causal frameworks. So, given some relatum $\mathcal{S}_k(t) = \bigcup f\mathcal{S}_j(t)$ for some j, we form the extension of the causal map as follows:

$$\mathfrak{C}(t_1, t_0) : \cup f\mathcal{S}_j(t_1) \longrightarrow \cup\mathfrak{C}(t_1, t_0)(f\mathcal{S}_j(t_1)) \tag{3.10}$$

The first union is in the causal framework at t_1 and the second union is in the causal framework at t_0.

Again we need not be too concerned, depending on context, as to whether we consider the right hand side as a set of fundamental relata or as a single relatum or a set of relata.

The causal pre-bundle together with its set of causal maps constitutes a **causal bundle**. Recall, the causal pre-bundle was the ordered collection of universe moments before we placed any connexion structure between the moments.

It is important to emphasize that the causal maps derive from the evolution of the universe, they characterize this evolution at the local level of fundamental causal relata. We are given the universe and its evolution and we then attempt to take this evolution to pieces in as much detail as is possible - that is what the causal maps do.

As a final point, we mentioned above that we expect the underlying manifold - as giving, perhaps, the synchronous spatial relationships between fundamental relata - to arise from the causal process rather than being imposed separately from it. This is not a crucial point in much of the arguments that follow. Whilst it was introduced as an assumption motivated by issues in contemporary physics, it also serves to ensure that our definition of a universe below is complete. If we allowed relata to exist as imposed on parts of some separate manifold then we could countenance areas of space that are devoid of causal activity. This causal emptiness, however, stands in contradiction to current physical thinking - such as the idea of the vacuum in quantum field theory or even issues surrounding the interpretation of gravitational and electric fields. We do not, however, attempt to argue for our assumption that fundamental relatum juxtaposition is derived from the causal process here as, whilst it may well be plausible from the way we have constructed our model, it is a contentious proposal. For instance, it could be argued that the underlying manifold is separate and both fully covered by relata - so there are no causally empty patches - and that the causal mechanism derives its structure, at least in part, from the juxtaposition of relata with respect to the manifold. This might then resolve some of the objections above but it does still leave issues connected with non-locality in quantum theory,[23] and issues with energy in general relativity.[24] We do not pursue these issues.

3.4 Causal maps and causal transitivity

In this section we introduce a further requirement on causality: transitivity. This will play a crucial, but largely uncommented, role in much of what follows. The conditions we present are motivated by the usual idea of causal transitivity: namely if event A causes event B and event B then causes event C then we can say that A was also a prior cause of C.

In passing, we mention there are some general conditions we should expect of the causal map:

$$\mathfrak{C}(\mathcal{S}_i \cup \mathcal{S}_j) = \mathfrak{C}(\mathcal{S}_i) \cup \mathfrak{C}(\mathcal{S}_j) \qquad \mathfrak{C}(\mathcal{S}_i \cap \mathcal{S}_j) = \mathfrak{C}(\mathcal{S}_i) \cap \mathfrak{C}(\mathcal{S}_j) \tag{3.11}$$

This is roughly to say that the cause of the combination of two events is the combination of their causes and similarly the cause of what is common to two events is common to the cause of those two events.

Now, transitivity: we shall require that the causal relation is transitive so that causes combine nicely across parameter values. We also assume that there are causal maps applicable to any relata for every pair of times other than at times that precede the occurrence of spontaneous relata in a causal history - something we deal with below. If we were not to impose this time-pair condition, we could have causality that is temporally non-local - where a relatum at one time caused a later relatum without this cause passing through intermediate moments. Using $*$ to indicate the combinations of causal maps, we have a condition for causal transitivity. Causal maps applied to some

[23] See, for instance, Bohm & Hiley (1993) or for an unpacking of non-locality, Redhead (1989).
[24] See Weyl (1999, §33, pp. 268-273) or Misner et al. (1973, pp. 466-468).

relatum, $f\mathcal{S}$, are $f\mathcal{S}$-transitive if they obey the following relation with respect to the relatum $f\mathcal{S}$:

$$\mathfrak{C}(t_2, t_1) * \mathfrak{C}(t_3, t_2) = \mathfrak{C}(t_3, t_1) \quad t_3 > t_2 > t_1 \tag{3.12}$$

And we speak of a universe, or causal bundle, being transitive if all the causal maps are transitive with respect to every relatum. Henceforth we assume that all causal maps are transitive with respect to all relata.

There might seem to be an issue with causal transitivity when the spontaneous coming into being of relata is countenanced. In such cases, any future relata caused by the actions of spontaneous relata do not have complete causes at times prior to the spontaneous relatum's occurrence. However, any worries dissolve by noting that in such cases the causal map, when applied to a relatum with a spontaneous relatum in its history, is only defined for that relatum at times at or after the occurrence of the spontaneous relatum in its history.

For the rest of this work we shall assume, without much further comment, there is no possibility of temporally non-local causality and we shall assume all universes are transitive.[25]

3.5 Introducing the measure

In this section we introduce some idea of quantification onto our structure. The ideas we introduce here will then play an important role in many of the arguments we set out in following chapters.

At each moment the full set of [structured] fundamental causal relata make up the universe at that moment. We assume that each fundamental relatum in these causal frameworks must have ontic causal content. That is, each fundamental relatum causally contributes to the content of subsequent relata through the connexion and this causal effect is derived from the content of the relatum in some manner - a manner that is left open here. If we were to talk in terms of contemporary physics we might say fundamental relata must possess content that builds particles and energy and momentum and buses and plants; and if we were to step beyond this then also thoughts and such like. **Ontic causal content**, or just **content**, is a primitive term here as it is to characterize any relatum that participates causally: later, in section 5.5, we shall put the case that it it plausible to assume that this encompasses aspects of 'the mental' too. What the content is may well be a mystery but that there is actual content sits as a basic doctrine of the causality we advocate here.

In this section we set up a measure which serves to quantify the content of any fundamental relatum and by extension any relatum. We shall then use this measure to place further restrictions on causal relations. There are theories of causality that have many similarities with aspects of what

[25] For some further discussion see, for instance, Schaffer (1995, §5).

we propose in this section. For instance the approach developed by Fair (1979) and that developed by Dowe (2000). Further, Kim (2007) seems to hint at a conception of causality that involves the transfer, or perhaps persistence, of actual content from moment to moment; although he does skew it towards the classically physical. Kim (2007) says,

> I am ... asking the reader to think of causation in terms of actual productive-generative mechanisms involving energy flow, momentum transfer and the like, and not merely in terms of counterfactual dependencies. Needless to say, the over-determination idea makes little sense when causation is understood this way. Kim (2007, n12, p. 47)

We shall see later that Kim is wrong to dismiss over-determination in the light of energy flow and so forth. Over-determination can only be dismissed if the quantity that is transferred by a connexion is also conserved and, additionally, no spontaneous creation of relata occurs.

The causal frameworks were set up to have the structure of a σ-algebra.[26] This allows us to place a measure on the structure. A measure is effectively a way of quantifying some aspect and here we take the measure to quantify the ontic causal content of relata. We are making the assumption that the measure we define is applicable across any framework in our bundle, so that whilst the structure and nature of the relata may well be different from moment to moment they all have measurable content. The implications of this assumption are touched on below.

The measure, μ, acts as a map from fundamental causal relata to positive real numbers. The number associated with a relatum by this measure is a quantification of its causal content. As causal content will be fundamental to the causal process and is taken as a the mark of existence then no relatum can have zero content.[27] We make a technical exception for the empty relatum ϕ. So, we have the measure as a map:

$$\mu : f\mathcal{S}_j(t) \longrightarrow \mu(f\mathcal{S}_j(t)) \in \mathbb{R}^+ \tag{3.13}$$

And this extends to causal relata that are constituted of any countable union of these fundamental relata:[28]

$$\mu : \mathcal{S}_j(t) \longrightarrow \mu(\mathcal{S}_j(t)) \in \mathbb{R}^+ \tag{3.14}$$

[26] For instance, see Halmos (1974).

[27] This does mean we cannot think of absence as having causal force. For a discussion, see Beebee (2004) or Schaffer (1995, §2).

[28] For a rigorous treatment, again see Halmos (1974).

Recalling that we defined fundamental relata so that they had no ontic content in common - their intersections are empty - then we see (3.14) derives from:

$$\mu(\cup f\mathcal{S}_j(t)) = \Sigma\mu(f\mathcal{S}_j(t)) \tag{3.15}$$

We also have:

$$\mu(\mathcal{S}_i(t) \cup \mathcal{S}_j(t)) = \mu(\mathcal{S}_i(t)) + \mu(\mathcal{S}_j(t)) - \mu(\mathcal{S}_i(t) \cap \mathcal{S}_j(t)) \; and \; \mu(\emptyset) = 0 \tag{3.16}$$

There is a slight issue with the universe as a whole but as we have assumed the universe is constituted of a finite or countably infinite union of fundamental relata then we can set the measure of the content of the universe to be finite at every time, even though it is consistent with what we have set up so far that it may vary from moment to moment.[29] Below we shall say more on the possibility of variation in what the measure quantifies across the bundle.

We can then define the contribution one fundamental relatum makes to another temporally juxtaposed relatum under the causal relationship; this then encapsulates the idea of transfer or influence that we mentioned earlier. We use the same symbol, μ, in both cases; the contribution of the earlier $f\mathcal{S}_i(t_0)$ to the later $f\mathcal{S}_j(t_1)$ is written as:

$$\mu(f\mathcal{S}_i(t_0), f\mathcal{S}_j(t_1)) \tag{3.17}$$

and we additionally impose:

$$\mu(f\mathcal{S}_i(t_0), \mathfrak{U}(t_a)) \leq \mu(f\mathcal{S}_i(t_0)) \;\; \forall t_a > t_0 \tag{3.18}$$

Where we interpret this as meaning no fundamental relatum can contribute more to any future content than it contains. We could relax this rule and allow a relatum to contribute its *total* content more than once to future states. There is no reason not to allow this but permitting such an increase of content suggests that something extra is created in the evolutionary process of the universe from moment to moment. Of course the plausibility of the assumption rests on what the measure quantifies. A steady-state theory may violate this measure restriction, especially if it measures something such as the matter or information content of relata.[30] Whilst we impose the

[29] We suggest that some motivation for associating a finite quantity with each relatum, and hence for the use of measures, is the Bekenstein bound, see Bekenstein (1981).

[30] For instance see Bondi & Gold (1948). However, there may be measures that are constant in any purported steady state theory - here we note that were we to allow the measure at each moment to be unrelated to that at other moments, we could increase the quantity of matter in the universe at each moment and then rescale any measure of it to be the same moment by moment.

restriction on our measure we must not be too dogmatic in the absence of a clear understanding of what the measure might quantify. The purpose of the measure, as we explain below, is to give some purchase to other causal concepts such as over-determination.

We said earlier that we shall assume that the content is of the same type at each moment but these contribution rules do not rely on this. We could merely require each moment to be measurable in some manner and for the universe at each moment to have a finite measure and then scale the total at each moment so that the total at each moment stays constant across the bundle. That is, we could abandon the idea that there is a common measure but keep a form of measure preservation artificially. However, this approach is rejected here both because it is not a fruitful approach and because we wish to assume that the universe is, in some sense, causally similar from moment to moment - in that there is a common notion of causality and causal content across all moments. We shall assume that the connexion between states requires that causal content be of the same sort at all times: content is measurable at all times by a single measure. Additionally, if we were not to do this then some of the definitions we derive below from the measure would be meaningless. Perhaps it would help to *think* of the measure as quantifying something familiar such as energy so that cause is driven by the temporal transfer of energy [or even information] in some manner. Any relatum that causes future states can contribute at most the energy it possesses and no more. That is to say the causal mechanism itself does not introduce increases in energy and any increase in energy would have to originate outside of the causal mechanism, for instance by spontaneous creation of relata.[31] As a way of *thinking* about the measure this should suffice for the moment. This allows us to characterize the cause at time t_0 of some given relatum, $\mathcal{S}$, at time t_1 as the set of *all* fundamental relata at t_0 that have a non-zero contribution to some fundamental causal relata that constitute $\mathcal{S}$.[32] We can write:

$$\mathfrak{C}(t_1, t_0)(\mathcal{S}_i(t_1)) = \{\mathcal{S}_j(t_0) : \mu(\mathcal{S}_j(t_0), \mathcal{S}_i(t_1)) > 0\} \tag{3.19}$$

With the exception of spontaneously occurring fundamental relata, we also assume the complete content of the effect derives entirely from its cause, so that, given:

$$\mathcal{S}_{cause} = \cup_i (f\mathcal{S}_i(t_0)) \ \ some\ i \tag{3.20}$$

[31]This is *not* to be thought of as the same thing necessarily as the spontaneous appearance of particles through quantum vacuum fluctuations, see for instance Srednicki (2007). This would most likely fall within the remit of causal relata already present at prior times.

[32]This allows for causal over-determination which we have yet to define.

and the corresponding effect:

$$\mathcal{S}_{effect} = \cup_j (f\mathcal{S}_j(t_1)) \ \ some\ j \tag{3.21}$$

then we have the ontic contribution of cause to effect as:

$$\mu(\mathcal{S}_{cause}, \mathcal{S}_{effect}) = \mu(\mathcal{S}_{effect}) \tag{3.22}$$

However, we need to be a little careful here. It would seem at first glance, and referring to (3.20) and (3.21), that we ought to define this equivalently as follows:

$$\mu(\mathcal{S}_{cause}, \mathcal{S}_{effect}) = \Sigma_i \Sigma_j \, \mu(f\mathcal{S}_i(t_0), f\mathcal{S}_j(t_1)) \tag{3.23}$$

This is incorrect as we have not yet taken account of the possibility of over-determination. At present, the best we can write is:

$$\mu(\mathcal{S}_{cause}, \mathcal{S}_{effect}) \leq \Sigma_i \Sigma_j \, \mu(f\mathcal{S}_i(t_0), f\mathcal{S}_j(t_1)) \tag{3.24}$$

We also assume that causal content is traceable through time and consistent with causal transitivity as defined above. This means that quantifications such as:

$$\mu(f\mathcal{S}_i(t_a), f\mathcal{S}_j(t_b)) \tag{3.25}$$

are meaningful and defined consistently. Whilst we choose to leave the exact definition of (3.25) unremarked as it might depend on exactly what the measure quantifies - information, energy-momentum etc - we note that the most natural [iterative] definition in simple situations is probably,

$$\mu(f\mathcal{S}_i(t_a), f\mathcal{S}_k(t_{n+1})) = \Sigma_j \left\{ \mu(f\mathcal{S}_i(t_a), f\mathcal{S}_j(t_n)) \times \frac{\mu(f\mathcal{S}_j(t_n), f\mathcal{S}_k(t_{n+1}))}{\mu(f\mathcal{S}_j(t_n))} \right\} \tag{3.26}$$

We do not pursue (3.26) any further here, preferring to assert merely that there is some well-defined way of unpacking (3.25) which is fully reflective of the ontic actuality.

This measure structure can now be used to characterize a number of common features of causality in a natural manner. It is to this we now turn.

3.6 Over-determination and epiphenomena

In this section we use the measure we introduced above to provide rigorous definitions of causal over-determination and epiphenomena. This will allow the clear use of both of these notions later and it will obviate extensive detours when they are mentioned there. We start by providing some general informal discussions of over-determination.

3.6.1 Prolegomenon to causal over-determination

In this section we seek to motivate, with some informality, the use of a measure to provide a rigorous definition of causal over-determination that we propose in the following sections.
Causal over-determination is often characterized as an effect having more than one sufficient cause. Two relata are said to causally over-determine some effect if each relatum individually would cause the effect in the absence of the other and the causing relata are independent and 'equal' with respect to their purported effect.[33] 'Equal' is taken to mean that both causal chains reach completion. That is to say, two relata over-determine their effect if they are both actual causes of the effect and the absence of one of the causal relata would not alter the occurrence of the effect in any way. There are three issues with this characterization that are relevant here: one is that it is not clear in what manner the absence of one cause is to be understood; second, there is an issue with event fragility; and third, characterizations of over-determination in standard instances are embedded in theory and dependent upon a particular conception of events and regularity. Now we turn to discuss these three issues in a little more detail.

The first issue is roughly to ask: what does it mean to claim that each event alone is sufficient to cause the effect? What are the criteria by which we are to judge the assertion as true? If we turn to possible worlds then we might hold that there are similar possible worlds to our actual world where one of the causes alone leads to the effect. We must then ask how similar this other world is to ours and, more particularly, by what criteria is this to be judged. Such questions are not easy to answer, especially when ontic concerns are brought into play - as we discuss in Chapter 4. If this possible world approach proves unsatisfactory as a route to an answer then we must ask how we are to grasp what is meant by the absence of one causing event in the scenario. Our first problem is then: how are we to unpack the claim that one event alone is sufficient to cause the effect in a consistent and perspicuous manner?

Now, the second issue. Consider the classic example of a firing squad executing a prisoner. The

[33] For some related discussion see Lewis (1987, pp. 159-213) and Schaffer (2003).

prisoner is hit simultaneously by two bullets and each effects a similar amount of damage. The prisoner dies. The usual claim is: because each bullet alone would have been sufficient to execute the prisoner, his death is over-determined by the two bullets. However the assertion that the death is over-determined requires an assumption to launch. We must adopt a generous appraisal of the event *death*. If we decide that *death*, or this prisoner's particular death, in this case is fragile then one bullet alone will cause a different death from the two bullets together.[34] The actual death caused by the two bullets is in detail quite specific and cannot be replicated by a single bullet acting alone. That is, if we take all macroscopic events as fragile instances then it is hard to maintain causal over-determination at the macroscopic scale.[35] So our second problem is to ask: what are we to mean by an over-determined event if its detailed structure is to remain entirely unaltered by the removal of an otiose cause?

The third issue is related to the first. We must worry that, allowing for a solution to the fragility question, the characterization of over-determination is embedded in theory.[36] The truth of the over-determination in the execution scenario requires the truth of the general fact that single bullets cause death. What if there could be no such fact? For instance, if the world is not amenable to a regularity where such facts are found. Instead maybe single bullets have myriad effects depending on the day of the week, the weather, and so forth; perhaps they turn their victims to living ectoplasm or sometimes have no effect whatsoever, passing through the target leaving no trace. If it were not the fact that single bullets, when identifiable, kill then over-determination could not be correctly characterized by stating that the absence of one event makes no difference to the effect.

It would seem that characterizations of over-determination might not be able to stand, even in principle, if the world is not actually regular in a form that is amenable to summary. If there is no regularity to the world at the level of causation then we cannot make sense of the multiplicity of sufficient causes by postulating that they can be removed without altering the effect's occurring unless the truth of the statement concerning removal is derived in a manner independent of regularity and theory. So our third problem is: how are we to understand over-determination in the absence of regularity and theory?

[34] Here we mean by 'fragile' that the exact event cannot occur in a different manner from the way it does occur or at a different time.

[35] For a more detailed exposition of this, see Bunzl (1979).

[36] As noted, this also applies to the possible worlds problem above as 'similarity' is usually derived from a notion of [epistemic] law-like behaviour which requires theory. See also section 4.4. We consider this third issue as distinct from the first as there may well be methods of eking out a notion of similarity of possible worlds without having to refer to regularity or theory.

The first issue is hard to resolve without appealing to possible worlds and/or counterfactuals. If some counterfactual approach is to be maintained and its unpacking using possible worlds rejected then some other understanding of counterfactuals would have to be advanced. And we put a case in Chapter 4 that counterfactuals are likely to clash with ontic concerns.

The second issue is very hard to resolve for individual events - every event is unique in its full set of specifiers - its place and time, the precise details of the atoms involved, which way the atoms moved and so on. Any resolution of this would have to appeal to a strong notion of repeatability within a universe or across universes.

The third issue is also difficult to defuse - in the absence of regularity there are no criteria for similarity across possible worlds and no recourse to theory to allow repeatability of 'identical' events occurring at differing spatio-temporal positions.

There are two broad ways out of these impasses: one is to argue carefully that they have no force and that the rather brief analysis presented here, when fleshed out, is mistaken or flawed. The second is to advance an alternative basis for over-determination which is sufficiently intuitively satisfying to be taken as a clear unpacking of the notion but which is free from the earlier ailments. This second approach is the one we shall take.

We shall define causal over-determination by making use of the measure we introduced in section 3.5. In the absence of such a measure this version of over-determination has no place. The version we advance has the added advantage that it allows for a gradation of over-determination, from the partial to the outright [pure].

3.6.2 Defining causal over-determination

Following the discussion above, we turn now to encapsulate causal over-determination formally. We distinguish in our definitions between relata that are causally over-determined and relata which are causally over-determining. We start with a definition of a relatum's being causally over-determined. Using (3.20) and (3.21) above, we have:

DEFINITION 3.4: Causally over-determined relatum

The relatum $\mathcal{S}_{effect}$ is **causally over-determined** when:

$$\mu(\mathcal{S}_{cause}, \mathcal{S}_{effect}) < \Sigma_i \Sigma_j\, \mu(f\mathcal{S}_i(t_0), f\mathcal{S}_j(t_1)) \tag{3.27}$$

This says that a relatum is causally over-determined if the total of the individual contributions from the fundamental relata that constitute its cause are considered separately and added then the result is greater than the total content the effect possesses. In such a case some of the contributions from the fundamental relata in the cause must be duplicated by contributions from other fundamental relata in the cause. We note here that this can only occur if contributions from individual fundamental relata are not strictly additive. This has implications: it might rule out maintaining both measure preservation and causal over-determination defined using the measure. It suggests that if the measure is, say, energy then we cannot maintain both causal over-determination and energy conservation unless we allow sufficiently [globally] coordinated energy creation elsewhere.

We now turn to define a causally over-determining fundamental relatum. First we note that (3.27) is not sufficiently fine-grained to allow us to pinpoint the source or sources of the over-determining contribution: it may come from one fundamental relatum or from some collection of them.
To remedy this in part, and to take over-determination to the level of individual fundamental relata, we propose to construct further more fine-grained definitions:

DEFINITION 3.5: Fully causally over-determining relatum

We say that one fundamental relatum, $f\mathcal{S}_i(t_0)$, is **fully causally over-determining** with respect to some given later fundamental relatum, $f\mathcal{S}_j(t_1)$, if the following two conditions holds:

$$f\mathcal{S}_i(t_0) \in \mathfrak{C}(f\mathcal{S}_j(t_1)) \tag{3.28}$$

$$\mu(\{\mathfrak{C}(f\mathcal{S}_j(t_1)) \mid f\mathcal{S}_i(t_0)\}, f\mathcal{S}_j(t_1)) = \mu(\mathfrak{C}(f\mathcal{S}_j(t_1)), f\mathcal{S}_j(t_1)) \tag{3.29}$$

Where $\mathfrak{C} = \mathfrak{C}(t_1, t_0)$ and $\{\mathfrak{C}(f\mathcal{S}_j(t_1)) \mid f\mathcal{S}_i(t_0)\}$ indicates the set of fundamental relata that make up the cause of $f\mathcal{S}_j(t_1)$ with $f\mathcal{S}_i(t_0)$ removed.

In **DEFINITION 3.5**, (3.28) says that $f\mathcal{S}_i(t_0)$ is part of the cause of $f\mathcal{S}_j(t_1)$. This ensures it does have a non-zero contribution. (3.29) tells us that removing $f\mathcal{S}_i(t_0)$ from the cause has no repercussions on the content of the effect. What this does not allow for is the possibility that the overdetermining contribution from $f\mathcal{S}_i(t_0)$ to $f\mathcal{S}_j(t_1)$ is less than the total causal contribution of $f\mathcal{S}_i(t_0)$ to $f\mathcal{S}_j(t_1)$. This motivates a further distinct definition of a partially causally over-determining relata as follows:

DEFINITION 3.6: Partially causally over-determining relatum

We say that one fundamental relatum, $f\mathcal{S}_i(t_0)$, is **partially causally over-determining** with respect to some given later fundamental relatum, $f\mathcal{S}_j(t_1)$ if the following three conditions hold:[37]

$$f\mathcal{S}_i(t_0) \in \mathfrak{C}(f\mathcal{S}_j(t_1)) \tag{3.30}$$

$$\mu(\{\mathfrak{C}(f\mathcal{S}_j(t_1)) \mid f\mathcal{S}_i(t_0)\}, f\mathcal{S}_j(t_1)) < \mu(\mathfrak{C}(f\mathcal{S}_j(t_1)), f\mathcal{S}_j(t_1)) \tag{3.31}$$

$$\mu(\mathfrak{C}(f\mathcal{S}_j(t_1)), f\mathcal{S}_j(t_1)) - \mu(\{\mathfrak{C}(f\mathcal{S}_j(t_1)) \mid f\mathcal{S}_i(t_0)\}, f\mathcal{S}_j(t_1)) < \mu(f\mathcal{S}_i(t_0)), f\mathcal{S}_j(t_1)) \tag{3.32}$$

Here (3.30) plays the same role as (3.28) above. (3.31) states that the contribution from $f\mathcal{S}_i(t_0)$ is not completely over-determined. And (3.32) states that there is some contribution that $f\mathcal{S}_i(t_0)$ makes to $f\mathcal{S}_j(t_1)$ which is over-determined.[38] When a relatum partially over-determines then it can, although it might not, also causally contribute to other relata in a non over-determining manner.

And finally, we set out to specify when a relatum does nothing other than causally over-determine some future relatum:

DEFINITION 3.7: Purely causally over-determining relatum

A fundamental relatum, $f\mathcal{S}_i(t_0)$, is **purely causally over-determining** with respect to some given relatum, $f\mathcal{S}_j(t_1)$, if it is both fully causally over-determining with respect to the relatum and, additionally:

$$\mu(f\mathcal{S}_i(t_0), f\mathcal{S}_j(t_1)) = \mu(f\mathcal{S}_i(t_0), \mathfrak{U}(t_1)) \tag{3.33}$$

Where $\mathfrak{U}(t_1)$ is the universe at time t_1. We could go on to explore the relationships between the causal transitivity condition outlined earlier and these set of definitions. We do not do this here.

[37] We could combine (3.31) and (3.32) into a double inequality. In the interests of clarity, we choose not to do so.

[38] Here we should note that whilst we are talking of content being partially over-determined, this is altogether different from asserting fundamental relata have causal parts, something we denied earlier. It may be the case that the measure of contribution quantifies the way one relatum influences another so that the whole content of a relatum participates in the influence but the amount it influences is less than the quantity of content it possesses.

3.6.3 Epiphenomenal relata

In simple terms, a relatum is epiphenomenal if it has a cause but does not then cause anything further. We can encapsulate this formally as follows:

DEFINITION 3.8: Epiphenomenal relatum

A relatum $\mathcal{S}(t)$ is **epiphenomenal** if it satisfies both

$$\mathfrak{C}(t, t_a)(\mathcal{S}(t)) \neq \emptyset \;\; some \;\; t_a < t \tag{3.34}$$

and

$$\mu(\mathcal{S}(t), \mathfrak{U}(t_b)) = 0 \;\; \forall t_b > t \tag{3.35}$$

Where (3.34) says that the relatum is caused and (3.35) says it has no further causal efficacy on the universe. We note that 'epiphenomenal' is sometimes used in another sense in the literature. For instance the term 'epiphenomenal' when applied to ectoplasm is used to mean 'causally isolated from the universe'.[39] There is an additional issue with this conception that we shall now turn, briefly, to deal with: we start by asking, 'what is the extent of a universe?'; that is, we ask what is it to posit the presence of elements within a universe that have no causal connection with any other relata there. In the next section we re-characterize a universe so that epiphenomenal ectoplasm cannot be considered as part of a universe. We do not argue against ectoplasm, rather we advance a definition that discounts its inclusion under the term 'universe'. We do this by insisting that to be counted as part of a universe any entity must have actual causal purchase on, or be caused by, something else that is already deemed to be within that universe. Epiphenomenal ectoplasm has, by definition, no causal purchase whatsoever and so it is not to be counted as part of a universe. We might speculate that the usual idea of epiphenomenal ectoplasm as being causally isolated requires a causally inert manifold in the background of any universe in question. Here we take it that the manifold of the universe, space-time or some such, arises from the causal structure so, for our models, epiphenomenal ectoplasm cannot have a location within a universe whilst remaining causally inert.

The relationship between the assumption of an underlying manifold on which causal mechanisms play out - as opposed to a manifold formed from the causal structures - and the possibility of epiphenomenal ectoplasm is an interesting direction to explore but we largely leave the matter

[39] See, for instance, Stoljar (2010, p. 134).

here.

3.7 Re-characterizing the universe's extent

In this section we move towards a characterization of what we are to mean by a **causally closed universe**, or **causal nexus** and, in passing, we also define again what is meant by a **deterministic universe** but by using some of the ideas we have developed in this chapter.
The aim here is to produce a clear delineation of what we are to mean by a 'universe' by characterizing all that it contains. The essential idea is that anything that either affects or is affected by elements that we consider to be within a universe is also then to be considered to be within the universe. We hold the only way in which elements within a universe can be affected is through causation.[40] That is to say that any element that cannot change or contribute to a relatum cannot be detected or have any force within this universe. This is a fundamental axiom of our causal model- and one we return to in section 5.5. It means, amongst other things, that we assume the mental is fully subsumed under causality - in that it must either be affected by or affect other relata in some manner.[41]

The structure we have proposed here allows for a universe that is causally closed but not necessarily deterministic. This is because our outline of causality does not come ready equipped with an idea of necessity. That is to say that, as our causal structure is derived from the actual evolution of a universe and does involve possibilities of predictability or probability, then relata have causal links even when the effects are not necessitated by the presence of the cause. For instance, in the double slit experiment we would hold that there is a causal link between one moment and the next throughout the process from releasing an electron through to its hitting the screen.[42] What we may not be able to say is that the corresponding relata at any one time uniquely determine the next relata but we hold that we *can* say that the relata at each point evolve into other relata and it is this evolution that gives us the causal map. If we were to deny that there is a connexion from moment to moment then it is not clear how any process of influence, probabilistic or deterministic, can be effected for fundamental relata. If fundamental relata must arise at each moment independently of those at previous moments, independently in that there is no content transfer or influence from moment to moment, then causality as we describe does not follow. That is, our Humean universes of Chapter 2 do not possess the notion of causality that we are outlining here.

[40]Later we discuss some issues that might arise were we to permit a certain form of supervenience into the mix. See sections 6.3.7 and 6.4.

[41]This is not an exclusive 'either . . . or'.

[42]For a discussion of the double-slit experiment, see for instance Feynman (2011, Vol III, Ch 1).

We start by defining a **causal nexus** generated by a fundamental causal relatum. For this we need to find every fundamental relatum that is related to our generator relatum using causal maps [backwards in time] and forward-maps which we shall define. The motivation behind this is to ensure that every fundamental relatum in the nexus also finds its cause *and* all those relata to which it causally contributes *both* sitting within the very same nexus. Removing any single relatum from the nexus then makes it incomplete and collapses the intricate causal relationships.

We define a **forward-map** for some given fundamental relatum at t_0, $f\mathcal{S}(t_0)$, via the following equation:

$$\mathfrak{F}(t_0, t_1)(f\mathcal{S}(t_0)) \equiv \{f\mathcal{S}_i(t_1) : \mu(f\mathcal{S}(t_0), f\mathcal{S}_i(t_1)) > 0\} \tag{3.36}$$

That is, the forwards-map gives all the fundamental relata at time t_1 that receive some non-zero causal contribution from $f\mathcal{S}(t_0)$. This allows us to use all possible forward-maps and causal maps, which recall went backwards in time, to generate a causal nexus. The causal nexus is made up from the web of connexions and from the [structured] set of fundamental relata that can be reached from the generating fundamental relatum by a path made up of links between fundamental relata that are formed from non-zero content contributions between the relata and the web of connexions:

DEFINITION 3.9: Causal nexus

A causal nexus generated by some given relatum, $f\mathcal{S}(t)$ is made up of:

1. All those relata generated by all maps made up of the form $...\mathfrak{F}(t_b, t_c) * \mathfrak{C}(t_c, t_b) * \mathfrak{F}(t_b, t_c)...\mathfrak{C}(t_p, t_z) * \mathfrak{C}(t_n, t_p) * ...\mathfrak{F}(t_q, t_r)...$ acting on the generating relatum. These are considered as forming structured sets at each time. That is, all possible allowable combinations of forward and causal maps for all possible time pairs.[43]

2. The network of connexions relating the structured sets of relata across time.

We note that any fundamental relatum in a causal nexus will generate the complete nexus via the process outlined in **1** of **DEFINITION 3.9**. We then characterize a universe as a [closed] causal nexus. It can be generated by any fundamental relatum that is within the universe given the set of causal and forwards-maps.

It may be that our universe, as a causal bundle, actually divides into a set of disjoint causal nexuses. However, recall that we hold that all 'mental' occurrences are themselves subject to our causal model.[44] We hold this because we assume all mental occurrences influence and/or are influenced by their surroundings - recalling we hold to a realist conception - and the only means of

[43]Recalling the restriction on the existence of causal maps resulting from spontaneous relata.

[44]And recall, we offer some discussions to justify that this is a *plausible* assumption in section 5.5.

diachronic influence we allow in our models is that of causality.[45] As the mental is subsumed as part of the causal structures of any universe in our model then we cannot have epistemic access to other disconnected parts of the causal bundle. This is because epistemic access is assumed to be underpinned solely by causal mechanisms. There is, by very definition, no experiment or method that we can use to find these other nexuses as all experiments require causal connexions. The characterization is set up in such a way that even relata that arise at some moment spontaneously, uncaused, and then go on to have causal consequences will be counted as part of a causal nexus; similarly for epiphenomenal relata. So on the assumption that we live in a causal universe, one where there are ontic connexions between moments, we have the epistemically accessible part of our universe forming part of the closed causal nexus. So for the rest of this study we shall take all connexion-bound universes to be causally closed nexuses. By contrast, a Humean universe is complete as a specification of states alone - a Humean universe is akin to a causal pre-bundle, a concatenation of universe states without any connexion between them.

It is important to realise that being causally closed is not the same thing as being deterministic. We can see this simply as a causal nexus permits the spontaneous, i.e. uncaused, formation of causally efficacious relata whereas, under all definitions, determinism does not.

We define a deterministic universe:

> **DEFINITION 3.10: Deterministic universe - causal characterization**
>
> A deterministic universe is a causal nexus where the universe at any given time [ontically] necessitates the states future to it.

This is meant to be identical in spirit to the definition advanced in Chapter 2. The only difference is that we have now set out firmly what we are to mean by a universe. Henceforth, when we talk of a deterministic universe we shall mean one that is both deterministic *and* a closed causal nexus. We require here that any deterministic universe is causally transitive. We mentioned earlier that we assume all universes are completely causally transitive. Recall we imposed this to avoid temporal non-local causation. Also it is useful to note that a deterministic universe precludes spontaneous relata and so all relata in deterministic universes have causal maps for every pair of times. We mention transitivity in the definition to allow that the universe at one moment necessitates all its future states. Abandoning transitivity makes the definition more complicated and we would have to allow for temporal non-local causality and thence require the whole past history, or some part thereof, to be what necessitates the future of the universe.

[45] This is not entirely correct: in Chapter 6 we do make some comments on the possibility of a synchronic ontic relationship between relata when we examine if the possibility of supervenience has any relevance to our arguments.

3.8 Causality and determinism

In this section we briefly explain why the models we have set up differentiate clearly between causality and determinism.

Causality is a relationship between relata mediated by a connexion and may or may not be an ontically necessary relationship. All that is required to state legitimately that one set of relata causes some later relatum is that there is a clear connexion, or influence, between them. This influence does not have to be one that is powered by necessity. As such, it is sometimes only possible to tell what caused a specific relatum in retrospect - we must take our ontic magnifying glass and examine the relata and how they relate via connexions. Once we have done so, we can tell which earlier relata have caused the later relatum. That is to say, in indeterministic cases, there is no Archimedean perspective that can tell what some given relatum will evolve into by merely looking at it. Laplace's demon has to wait to see what happens and then reconstruct causality from the connexions that are formed as the universe evolves.

Determinism, on the other hand, requires both the presence of a connexion *and* that this connexion relates relata necessarily. Looking with an ontic magnifying glass, and armed with an Archimedean perspective, will allow us to tell what will happen in the future. In this case the evolution and the nexus of connexions can be anticipated in advance by Laplace's demon.

Causality is our name for, loosely, the relationship between relata across time - a relationship that must be mediated by a connexion. Determinism is more than this - it requires ontic necessity to accompany the relationship.

Any attempt to link causality and determinism must be invalid within our model - and we take it as invalid and confused in general as it muddles issues that are seen here to be quite separate.

In other words, we can have the causal map:

$$\mathfrak{C}(t_1, t_0) : f\mathcal{S}_i(t_1) \longrightarrow \{f\mathcal{S}_j(t_0);\ some\ j\,\} \tag{3.37}$$

but we cannot deduce from this any ontic necessity:

$$(\mathfrak{C}(t_1, t_0) : f\mathcal{S}_i(t_1) \longrightarrow \{f\mathcal{S}_j(t_0);\ some\ j\,\}) \nRightarrow (\{f\mathcal{S}_j(t_0);\ some\ j\,\} \mathrel{\boxdot\!\!\rightarrow} f\mathcal{S}_i(t_1)) \tag{3.38}$$

It might then seem tempting to set the scope of ontic necessity for relata using:

$$(\{f\mathcal{S}_j(t_0);\ some\ j\,\} \mathrel{\boxdot\!\!\rightarrow} f\mathcal{S}_i(t_1)) \Rightarrow (\mathfrak{C}(t_1, t_0) : f\mathcal{S}_i(t_1) \longrightarrow \{f\mathcal{S}_j(t_0);\ some\ j\,\}) \tag{3.39}$$

But we must be careful. Expression (3.39) is only valid if there is no causal over-determination. If there were over-determination then a subset of the cause of $f\mathcal{S}_i(t_1)$ would be sufficient to ontically necessitate $f\mathcal{S}_i(t_1)$ and so $\{f\mathcal{S}_j(t_0);\ \textit{some } j\}$ may not be the compete cause of $f\mathcal{S}_i(t_1)$ as it need not include some of the over-determining relata.

If we assume there is no causal over-determination, then we can assert (3.39).

3.9 Separability, identity and essentiality

In this section we use our model above to help set out formal definitions of identity, essentiality and separability for both universes and [fundamental] relata. In simple terms, essentiality is the idea that the 'things' in a universe carry their nature inherently - as opposed to having it imposed from the outside. Separability is, again in very simple terms, the opposite of essentiality in deterministic settings - it holds that the 'things' of the universe gain their causal nature from a source separate from themselves.

The definitions we construct here will then be drawn upon in some of the arguments in following chapters and in section 3.10.

In section 3.10, we demonstrate that the definition of universe separability we offer in this section does not sit comfortably with our definition of ontic determinism. We also explain that this does not then lead directly to essentialism in its usual guise - such as in Ellis (2007); rather, we show there is an intermediate form of essentialism - less restrictive than Ellis's version - which applies to deterministic universes as a whole but does not reduce to essentiality of their constituting relata and so does allow some local separability. Ellis's essentialism is based on both the existence of natural kinds and regular laws. His idea is closest to our notion of relatum essentiality as set out in **DEFINITION 3.18** below. He does not allow for our broader version of universe essentiality.

First, identity. We ask, what are the criteria for two universes, as defined primarily in section 3.7, to be identical and we also ask what are the criteria for two relata to be identical.

For universes we can set up two distinct definitions of identity, one based on states alone and one based upon states considered together with the relationships between the states - the web of connexions.

Recall, from our definition above, a universe is a collection of time ordered moments that are additionally related from moment to moment by a web of influences, or connexions. Denoting a universe by $\mathfrak{U}_i$, its states by $\mathfrak{U}_i(t)$. Time is discretely indexed, so that all the time values come from a set where time is discretely parametrized: $t \in \mathcal{I}$; and denoting the relevant connexions by $\mathfrak{C}_i$, we can denote a connexion-bound universe by:

$$\mathfrak{U}_i = \{\mathfrak{C}_i; \mathfrak{U}_i(t), \ \ t \in \mathcal{I}\}$$

And, similarly in an obvious way, a Humean universe by:

$$\mathfrak{H}_i = \{\mathfrak{H}_i(t), \ \ t \in \mathcal{I};\}$$

We can also consider a universe up to and including some time, t_a, by writing:

$$\mathfrak{U}_i(t \leq t_a) = \{\mathfrak{C}_i(t \leq t_a); \mathfrak{U}_i(t), \ \ \forall t \in \mathcal{I}, t \leq t_a\} \tag{3.40}$$

And we could construct a similar expression for Humean universes:

$$\mathfrak{H}_i(t \leq t_a) = \{\mathfrak{H}_i(t), \ \ \forall t \in \mathcal{I}, t \leq t_a\}$$

We are now in a position to define when two universes are identical. We assume that the identity of two states of a universe is clearly defined but we make some comments on this below; and we also assume that it is clear when two connexions between states are identical but we shall make some comments below. We split up the idea of identity into two subtypes: state-identity and complete-identity. We denote both connexion-bound and Humean universes by $\mathfrak{U}$ when they appear in our definitions:

DEFINITION 3.11: : State-identical universes
Two universes, $\mathfrak{U}_i$ and $\mathfrak{U}_j$ are **state-identical**, denoted $\approx$, if and only if $\mathfrak{U}_i(t) = \mathfrak{U}_j(t)\ \forall t$

Where the '=' in $\mathfrak{U}_i(t) = \mathfrak{U}_j(t)$ is taken to mean they are constituted of identical sets of fundamental relata - they have the same elements - and these sets are structured in exactly the same way.

And complete-identity, which does not apply to Humean universes:

DEFINITION 3.12: Complete-identical universes
Two connexion-bound universes, $\mathfrak{U}_i$ and $\mathfrak{U}_j$ are **complete-identical**, denoted $\cong$, if and only if they are both state-identical and $\mathfrak{C}_i = \mathfrak{C}_j$.

Noting:

$$\mathfrak{U}_i \cong \mathfrak{U}_j \Rightarrow \mathfrak{U}_i \approx \mathfrak{U}_j \tag{3.41}$$

But,

$$\mathfrak{U}_i \approx \mathfrak{U}_j \nRightarrow \mathfrak{U}_i \cong \mathfrak{U}_j \tag{3.42}$$

We also propose that the identity relations extend to universes specified up to some time. So we have:

$$\mathfrak{U}_i(t \leq t_a) \approx \mathfrak{U}_j(t \leq t_a) \Leftrightarrow \mathfrak{U}_i(t) = \mathfrak{U}_j(t) \;\; \forall t \leq t_a \tag{3.43}$$

and

$$\mathfrak{U}_i(t \leq t_a) \cong \mathfrak{U}_j(t \leq t_a) \Leftrightarrow (\mathfrak{U}_i(t) = \mathfrak{U}_j(t),\; \forall t \leq t_a) \wedge (\mathfrak{C}_i(t \leq t_a) = \mathfrak{C}_j(t \leq t_a)) \tag{3.44}$$

Additionally it is useful to notice, that if we are given two connexion-bound deterministic universes up to some time, t_a, $\mathfrak{U}_i(t \leq t_a)$ and $\mathfrak{U}_j(t \leq t_a)$, we have:

$$\mathfrak{U}_i(t \leq t_a) \cong \mathfrak{U}_j(t \leq t_a) \Rightarrow \mathfrak{U}_i \cong \mathfrak{U}_j \tag{3.45}$$

but later we shall see that this relationship serves to trouble the idea of separability we outline below. We should also note:

$$\mathfrak{U}_i(t \leq t_a) \approx \mathfrak{U}_j(t \leq t_a) \nRightarrow \mathfrak{U}_i \approx \mathfrak{U}_j \tag{3.46}$$

We now turn to define when two fundamental relata are identical and when two relata are identical.

DEFINITION 3.13: Content-identical fundamental relata

Two fundamental relata are **content-identical**, denoted $\equiv$, if they have identical causal content.

Where we assume, as previously discussed, the causal content of a relatum is primitive. This allows us to move to define when two relata are identical. We have:

DEFINITION 3.14: Content-identical relata

Two relata $\mathcal{S}_i = \{f\mathcal{S}_{i1}, f\mathcal{S}_{i2}, ..., f\mathcal{S}_{in}\}$ and $\mathcal{S}_j = \{f\mathcal{S}_{j1}, f\mathcal{S}_{j2}, ..., f\mathcal{S}_{jn}\}$ are **content-identical** if and only if m = n and also $(\forall f\mathcal{S}_i \in \mathcal{S}_i, \exists f\mathcal{S}_j \in \mathcal{S}_j : f\mathcal{S}_i \equiv f\mathcal{S}_j)$ and $(\forall f\mathcal{S}_j \in \mathcal{S}_j, \exists f\mathcal{S}_i \in \mathcal{S}_i : f\mathcal{S}_j \equiv f\mathcal{S}_i)$ and the sets are structured in the same way.

In other words, relata are content-identical if they are composed of suitably defined [structured] sets of fundamental relata that are related through content-identity - including possible duplicates of fundamental relata if we allow content-identical fundamental relata to occur more than once in any given relatum.[46]

[46]We leave the term 'same' to refer to the very same relatum. So to say that $\mathcal{S}_i$ is the same as the relatum $\mathcal{S}_j$ means we have just labelled one single relatum in two different ways - and perhaps discovered this fact in some other way.

We now turn to define separability formally. Separability for universes is the name we give to the idea that the ontic laws and the states somehow arise, at least in part, separately from each other. And recall that separability can only be applied in deterministic settings as indeterministic universes [and indeterministic relata] although possibly connexion-bound, do not have ontic laws governing their evolution. The idea of separability might seem useful at it seems to be what is needed to allow us to talk of a concatenation of states, separately from a universe - i.e. a concatenation of states together with a connexion. Why might this be useful? It is useful to do this if we want to talk about possible worlds that mimic ours in some manner but not in others. For instance, we might wish to talk of worlds that are similar to ours up to some time - in that they have the same configuration of objects up to that time - and then diverge from our world in some [small] manner. Such scenarios occur when attempts are made to examine counterfactuals using possible worlds - something we return to in section 4.4 of Chapter 4. In other words, separability is the ontic notion that seems to underpin the idea that we can talk of tweaking and changing things in our universe a little without worrying about wider repercussions.

Equally, we have the idea that ontic laws are essential to a universe when they arise directly from the way the state of the universe is configured and its past history. We say that a deterministic universe that is driven forwards by essential laws is an essential universe. We define this as follows:

DEFINITION 3.15: Essential universe

A deterministic universe $\mathfrak{U}_i$ is **essential**, if given any universe $\mathfrak{U}_x$ such that $\mathfrak{U}_i \approx \mathfrak{U}_x$ then it follows that $\mathfrak{U}_i \cong \mathfrak{U}_x$.

And this allows us to define a separable universe:[47]

DEFINITION 3.16: Separable universe

A deterministic universe $\mathfrak{U}_i$ is **separable**, if given any universe $\mathfrak{U}_x$ such that $\mathfrak{U}_i \approx \mathfrak{U}_x$ then it does not follows that $\mathfrak{U}_i \cong \mathfrak{U}_x$.

We can also define separability for fundamental relata, and by extension for relata:

DEFINITION 3.17: Essential fundamental relatum

A fundamental relatum, $f\mathcal{S}$ is essential if it ontically necessitates its contribution to future states and this contribution depends solely on the relatum's content. Its contribution is independent of both the time of occurrence of $f\mathcal{S}$ and the position,[48] [or place in universe's configuration of relata] of $f\mathcal{S}$ relative to other relata.

[47] Below, in section 3.10, we shall discuss how this sits in tension with equation 3.45.

[48] By this we do not mean to say spatial position necessarily. But we must have spatial structures arising from, or present prior to, relata evolution so there must be some larger structural relations between relata that are related in some manner to a manifold, and thence in some manner to space. The idea of position is inherited from this manifold. The discussion on relata and objects in section 6.3.2 is also apposite to this issue of space and relata.

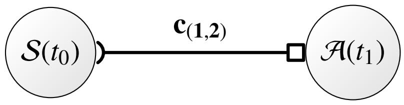

Figure 3.1: Illustration of an essential relatum

This says, informally, that a fundamental relatum is essential if it must be the way it is causally and cannot be otherwise. The implication is that the full nature of the way it impinges causally on the future comes from the way it is constituted in itself. We could have given the definition in terms of the behaviours of content-identical relata elsewhere but we chose not to as we do not want to anticipate that any fundamental relatum occurs more than once in any universe. Of course, by extension, if a fundamental relatum is essential and has content-identical copies, then these will all impinge causally on the future in the same way.
By extension we can say:

> **DEFINITION 3.18: Essential relatum**
>
> A relatum is **essential** if it is constituted of a [structured] set of fundamental relata that are all essential.

We illustrate an essential relatum in Figure 3.1. The rounded parenthesis on the arrow-end denoting the essential nature and the square box denoting ontic necessity. Of course, $\mathcal{A}(t_1)$ is not necessarily a compete set of fundamental relata - as the discussion on a causal penumbra above, in section 3.3, clarifies.

We can then build similar definition of a separable fundamental relata:

> **DEFINITION 3.19: Separable fundamental relatum**
>
> A fundamental relatum, $f\mathcal{S}$ is **separable** if its contribution to future relata is ontically necessitated and it is not an essential fundamental relatum.

Informally, this says that to be separable a relatum must ontically necessitate its future relata but the manner in which it does so, and hence the way it impinges causally on the future, can vary and so is independent of the relatum's causal content.

We extend the definition naturally to relata, with the same caveat as before:

> **DEFINITION 3.20: Separable relatum**
>
> A relatum is **separable** if it is constituted of a set of fundamental relata that are all separable.

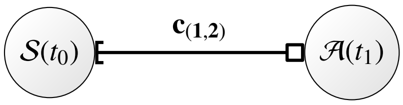

Figure 3.2: Illustration of a separable relatum

We illustrate a separable relatum in Figure 3.2. We do not pursue hybrid cases here.
We can now clarify the relationship between essential and separable relata and essential and separable universes. It is not as straightforward as might be expected. It is worth recalling in the following discussion that separability and essentiality are only applicable to deterministic settings.

First, if a universe is separable then it need not be the case that all its relata are separable too. It is quite possible to maintain the definition of universe separability with some, but not all, of the constituting relata being essential. This is because separability for a universe merely means that the ontic laws are not completely fixed for that universe's states,[49] it does not extend to say that there is necessarily complete freedom on how ontic laws may be modified in the transition to other possible worlds [universes]. We could, of course, modify the definition of a separable universe to make it true that all its relata must also be separable:

> **DEFINITION 3.21: Locally-separable universe**
>
> A universe is **locally-separable** if and only if all its relata are separable.

Equally it is not the case that an essential universe must be made up of essential relata. This is because it is possible that the essential nature of the ontic laws for a universe do not reduce to essential behaviour at the local level - the level of relata. It could be the case that the ontic laws are essential but dependent upon the configuration of the whole universe at each moment. This would then allow content-identical relata to behave differently at different times, and hence allow for the possibility that some relata are separable even though the universe itself is essential.
Again, we might strengthen our definition:

> **DEFINITION 3.22: Locally-essential universe**
>
> A universe is **locally-essential** if and only if all its relata are essential.

Let us summarise the possibilities we have here:

> **[i]** separable universe $\nRightarrow$ separable relata
>
> **[ii]** separable relata $\nRightarrow$ separable universe
>
> **[iii]** locally separable universe $\Leftrightarrow$ separable relata

[49] It is for this reason that universe separability clashes with determinism. We explain this below in section 3.10.

[iv] essential universe $\not\Rightarrow$ essential relata

[v] essential relata $\Rightarrow$ essential universe

[vi] locally essential universe $\Leftrightarrow$ essential relata

Noting that **[iv]** explains **[ii]**.

3.10 Separability and determinism

In this section we argue that a deterministic universe cannot be separable. That is, we argue the definitions of universe separability and determinism clash. By **[i]** above this does not then allow us to deduce that the relata that make up a deterministic universe are not separable. We shall have some more to say on relata separability in the next chapter.
What the arguments here, and their continuation in Chapter 4, amount to is this: we cannot modify deterministic aspects of any universe locally - affecting only single relata at a given time - whilst remaining true to the ontic features of the universe.

Recall that we defined a deterministic universe in section 2.3 of Chapter 2 as follows:

DEFINITION 3.23 Deterministic universe - alternative definition
A connexion-bound universe $\mathfrak{U}(T)$ is **deterministic** if and only if $\forall \mathfrak{U}(t_n) \in \mathfrak{U}(T)$ it is the case that $\mathfrak{U}(t_n) \boxdot\!\!\rightarrow \mathfrak{U}(t_{n+1})$.

We also held, in section 2.10.3, that for a universe to be deterministic its ontic laws must be set in advance and carried along with the universe as it evolves. Here we shall explain that there is some difficulty if we admit separable universes to the mix.
Consider a direct consequence of **DEFINITION 3.23**, namely what we stated for (3.45). There we said,

> ...given two connexion-bound deterministic universes up to some time, t_a, $\mathfrak{U}_i(t \leq t_a)$ and $\mathfrak{U}_j(t \leq t_a)$, we have: $\mathfrak{U}_i(t \leq t_a) \cong \mathfrak{U}_j(t \leq t_a) \Rightarrow \mathfrak{U}_i \cong \mathfrak{U}_j$.

However, there is a tension between this characterization and separability. First, consider the definition of separability we offered above; this said:

DEFINITION 3.16: Separable universe
A deterministic universe $\mathfrak{U}_i$ is **separable**, if given any universe $\mathfrak{U}_x$ such that $\mathfrak{U}_i \approx \mathfrak{U}_x$ then it does not follow that $\mathfrak{U}_i \cong \mathfrak{U}_x$.

That is, for a separable universe:

$$\mathfrak{U}_i \approx \mathfrak{U}_j \not\Rightarrow \mathfrak{U}_i \cong \mathfrak{U}_j$$

and we see this is in tension with the definition of determinism as we cannot move from $\mathfrak{U}_i(t \leq t_a) \approx \mathfrak{U}_j(t \leq t_a)$ to $\mathfrak{U}_i(t \leq t_a) \cong \mathfrak{U}_j(t \leq t_a)$ and hence to $\mathfrak{U}_i \cong \mathfrak{U}_j$. So it is clear that if two universes are separable they need not obey (3.45). What does this mean? It tells us that our definitions do not allow for separable universes to be deterministic. It need not be the case that the state of a separable universe up to some point ontically necessitates a unique future for that universe.
There is a good reason why this difficulty has arisen: it is because we have insisted that the ontic laws that drive the universe forward are set in advance *and* somehow inhere within the fabric of the universe - that is, they are present in the very fabric of the states in some manner. Recall, we discussed this in section 2.10.3 of Chapter 2. Universe separability allows us to separate out the ontic laws from the universe itself, so that they no longer inhere within it. But if we do hold to our definition of determinism and allow nothing from outside the sets of states to interfere in the evolutionary process then we cannot maintain the coherence of the idea of universe separability.

What are we to do? Should we modify our definitions of determinism to accommodate universe separability or maintain them and take it that separability is not a viable category for universes?

We can hold to the definition of determinism and also to universe separability but we have to abandon the idea, for separable universes, that the ontic rules are inherent in the universe's fabric; rather ontic laws must then be imposed from outside of the states and hence, by our definition of universe, from outside of the universe too. That is to say, to maintain the coherence of universe separability in the face of ontic determinism, we must postulate a set of laws that are set out across the sweep of time but which are not carried along in the fabric of universe states; instead these laws must somehow be accessed - non-causally - and acted upon at each moment by the states at that moment: the instructions as to how to proceed, how causes relate, must be set *outside* of the relata, *outside* of the universe, and imposed upon them.
It might be possible to maintain this - making laws external to the universe in some manner - but it would take us to a model that is more involved than ours. We would require both a temporally evolving set of states and, in addition, a very specific [timeless] realm in which our laws are to be deposited so that they might then impose their will on the states. And then we would have to modify, or add codicils to, our definitions to accommodate these additions.
The worry for us is that separability is often used when transiting across possible worlds: the

argument usually proceeds using worlds constructed to be very much like ours - populated by electrons, billiard balls, people, cheesecake and the like - but with different ontic laws, perhaps modified only locally, governing them. But, if we are to be able to justify this sort of similarity-preserving-transition and maintain an ontic construal of what possible worlds tell us then we must also permit the altering of ontic laws without substantially changing the nature of the universe itself. And to allow that is to place a schism between the content of the universe, as time-bound, and the laws as some existing actual 'thing' or repository outside of time and accessed with certain restrictions in place.

So, if we are to permit any meaningful talk of worlds like ours but with ontic laws of nature that differ we must be careful. We must require a clear separation of the content of the worlds - the states - and the manner in which they are forced to evolve from moment to moment - the ontic laws. And we must further ensure that this separation is coherent as a model.

Equally, when arguing using universes that are [largely] the same as ours except for some small change here or there in content we must be careful. If we wish to change the content by changing laws at some point then we can only do so on the assumption that the universe is separable - as otherwise to change a law is to change a relatum and to change a relatum is to change the whole past causal nexus and the causal relations that lead to this relatum.[50] Similarly, if we wish to add something to a state in the universe, such as an ammonia molecule,[51] whilst keeping those relata far away sufficiently similar to those in our own actual universe then we must be aware of the assumptions we must make. That is all to say, we must be cautious in changing the laws as we must ask how far can we stretch our claim that the new universe is similar to the one we inhabit - similar in ontic terms as opposed to epistemic terms. The point to make here is that the conceivability of similar universes to ours may not reflect reality when ontic restrictions are brought into play and some account must be made of these ontic restrictions if hidden assumptions are not to be smuggled into arguments.

In Chapter 4, we look in some detail at how the ontic nature of a universe restricts what we can plausibly change as we transit from it to other universes that we might be tempted to consider as 'similar'.

[50]We simplify matters here a little. We may be able, in some circumstances, to change a relatum without affecting past states: for instance, if the relatum in question has no causal history - so that it is spontaneously created. Of course, such relata can only occur in indeterministic universes and the question can then only relate to relata separability and not universe separability.

[51]The idea of adding an ammonia molecule to a universe was introduced by Kim, albeit when discussing supervenience - see Kim (1993, p. 85).

3.11 Taking stock

Let us pause to take stock of the important conclusions we have derived in this and the previous chapter. We have a mix of notions - determinisms, predictability, causality, separability, essentiality - which are often used implicitly or explicitly in debate. The main conclusions that we shall carry forward to the next few chapters stand as follows:

- Determinism is an ontic concept dependent upon ontic necessity;
- Ontic necessity does not give us predictability;
- Determinism does not necessarily give us predictability;
- We have epistemic access to the states of the universe;
- Epistemic access is not sufficient to determine the type of universe we inhabit;
- Indeterminism is compatible with predictability;
- Predictability tells us nothing as to the ontic nature of our universe;
- A deterministic universe cannot be separable;
- A deterministic universe can be locally-separable;
- A deterministic universe must be essential;
- A deterministic universe need not be locally-essential;
- Causality and determinism are separate doctrines.

And we note that Ellis' essentialism - as set out in Ellis (2007) - as being local essentialism, does not follow from determinism although our essentialism does. The difference is our essentialism only requires the ontic laws to be carried with the fabric of the universe as a whole and makes no demands on the local structure. Indeed, it is possible for all a universe's relata to be separable but the universe itself to be essential.

We are now in a position to move to the issues we tackle in the following chapter. We have put a skeleton case above that we must be very careful when attempting to construct a notion of similarity across possible worlds because any such notion would have to carry with it some ontic features and be set within ontic constraints. It seems that if such ontic features inhere in the world then a single attempted adjustment at one spatio-temporal location when passing to a similar world is in fact forbidden. This is because the parts of each world, at least those that are deterministic, are glued together with a web of connexions in such a manner that breaking and reassembling a single connexion may not be possible if the deterministic nature of the world is to be maintained or any notion of similarity is to unpack nicely. We take this on board in the next chapter where we argue, in part, that these ontic restrictions have repercussions on the way we understand the possibility of the 'choices' a universe may offer up.

3.12 Appendix: causal sets

In this section we show, albeit briefly, that there is a relationship between our causal structures and causal sets. We look at this briefly as we take this similarity as adding weight to the plausibility of the scheme we have proposed here but we do not make extensive comments on the similarities between the structures.[52]

We do not provide here a detailed motivation for causal sets. Suffice it to say that they are a mathematical scheme developed in response to issues that arise in trying to meld quantum theory and general relativity together. An excellent discussion of the background and rationale behind causal sets can be found in Dowker (2006). There she says,

> Causal set theory is an approach to the problem of quantum gravity in which spacetime is fundamentally discrete and in which causality is a primary concept. The concrete kinematics of the approach makes it relatively straightforward to build phenomenological models that produce observable effects of the underlying discreteness. (Dowker, 2006, Abstract, p. 1).

Here we intend only to show that, under some additional restrictions, the causal structure we developed above satisfies the mathematical axioms for being a causal set.

We start by quoting Dowker:

> Mathematically, a causal set is a locally finite partially ordered set, or in other words a set C together with a relation $\prec$, called 'precedes', which satisfy the following axioms:
> **(1)** if x$\prec$y and y$\prec$z then x$\prec$z, $\forall$ x, y, z $\in$ C (transitivity);
> **(2)** if x$\prec$y and y$\prec$x then x = y, $\forall$ x, y $\in$ C (noncircularity);
> **(3)** for any pair of fixed elements x and z of C, the set $\{y : x \prec y \prec z\}$ of elements lying between x and z is finite.
> (Dowker, 2006, p. 4)

To ensure our causal theory satisfies the axioms of causal sets we need to specify three things:

[CS1] $\mathfrak{C}(t, t) = I$ where I is the identity map sending a fundamental relatum to itself;

[CS2] $\mathfrak{C}(t_b, t_a)(f\mathcal{S}(t_b))$ is a finite set of fundamental relata $\forall t_a, t_b$;

[CS3] The total number of moments within a universe is finite.

[52] We do not, for instance, explore time in causal set theory.

We then consider the set of all fundamental relata partially ordered by the causal map to be our causal set:

For any two fundamental relata x and y we require: $x \prec y \Leftrightarrow \exists \mathfrak{C} : x \in \mathfrak{C}(y)$

Where $\mathfrak{C}$ is some causal map. From this, we have:

(1) is satisfied by virtue of (3.12)

(2) is satisfied by virtue of **[CS1]**;

(3) is satisfied by virtue of **[CS2]** and **[CS3]**.

Chapter 4

The ontic nature of choices

4.1 Introduction

Our aim in this chapter is to provide an ontically motivated understanding of what it is for a universe - indeterministic, deterministic or hybrid - to present genuine 'choices'. And by 'choices' we mean both that the universe presents, at some moment, ontically real alternative possible future paths and that each self, the conscious denizens of the world, can then select between them in a manner that makes them responsible for their choices. Recall, we said in Chapter 1 that we seek a model for free will which satisfies the two necessary conditions:

1. The universe presents genuine choices, genuine alternative futures, to its inhabitants;

2. The inhabitants 'choose' between these choices in such a manner that they are clearly and solely responsible for their choices, their selections.

In this chapter, we concern ourselves with **1** but we keep the term 'choices' in deference to what will follow in Chapter 5. We put aside the difficult issues of what it is for some being to make choices, to select between alternatives, or what it might be for a being to make free choices. That is all to say, here we are going to explore what features of a universe might allow us to claim genuine choices are presented by it, but without worrying about the way these choices are then selected and what we might require of beings for them to be 'free' in selecting between these choices.

4.2 Prolegomenon to the problem of ontic choice

In this section we set out one way of unpacking the issue we wish to address: that of providing an ontic construal of choices. We put our problem more formally in section 4.3 and then proceed to argue towards our conclusions in detail.

Intuitively we take choices to involve two aspects: one, that we are presented with alternative future paths that we might select between; and two, we do then select between these alternative paths in some manner. We usually take our selection to be intentional - we think things through, we weigh up the options, and then choose one path over the others and act upon our choice. In this chapter what is important is that the alternative futures are actually genuine and not merely derived from our conceptions of the world. And by 'genuine' we mean that they must have some [ontic] presence in the present - that is, there must be something in the present that indicates the alternative future paths are genuine possibilities for the future.

We might be tempted to think that a set of conceptual alternatives is adequate to provide choice. We might be inclined to think this because it might seem that it is within our power to choose between a set of conceptual alternatives and then pilot the world towards our choice. I can conceive that I might turn either right or left at an upcoming road junction. I might then choose to turn left and then pilot my actions towards this outcome turning my conception into an actuality whilst letting the other fizzle away unrealized. If this is the case then conceptual alternatives might provide what we seek: conceptual alternatives and the ability to pilot oneself towards one amongst them may encapsulate genuine possibilities for the future. However there are two reasons this is not a possible unpacking of choice.

First, there is no guarantee that conceptual alternatives reflect actual alternatives - especially if we inhabit a deterministic world. In this chapter we argue that deterministic universes are not open to alternative futures and also that, in the indeterministic case, the situation is not as simple as often assumed.

Second, we cannot argue that we have the power to choose between conceptual alternatives and then make our choices into an actual future through our actions. We cannot argue that conceptual alternatives are real alternatives because we have the power to make any of them happen. We cannot argue for this possibility because, as we show in sections 5.6 and 5.7, we do *not* have the power to pilot our actions in such a way that they could move from conceptual alternatives to a single actual future. And further, if conceptual alternatives are to be reflective of genuine possibilities they *must* all be possibilities for the universe when they are conceived - and it is this that, as we shall argue in this chapter, neither determinism nor pure indeterminism permits.

So, in informal terms, we are seeking some ontic understanding of what it is for a universe to offer

alternatives - choices - at moments in its history. We begin by outlining why this is not a trivial issue that can be solved by merely nodding towards indeterminism as it stands.
First, three observations:

[DET] In a deterministic universe everything in the future is made ontically necessary by the present. Recall this means that the fabric of the present moment is sufficient to ontically necessitate a unique future. And recall further that we explained at the end of Chapter 3, in section 3.9, that the ontic laws are themselves essential to the universe and inherent in the very fabric of the set of relata from which it is made. If we were to take up an Archimedean position and use our ontic magnifying glass to examine this fabric we should find everything we need there to tell us in exacting detail what the future states will be. That is to say, the future states have an ontic presence in the current moment: they are enfolded within it.

[INDET] If we were to examine an indeterministic universe we would find no embryonic future enfolded in its fabric. For simplicity here we shall consider a universe that contains no essential relata and no possibility of ontic necessity between relata - we call this 'pure indeterminism'. In such a universe the future is not fixed until it occurs. There is nothing in advance of its occurrence that can be taken to indicate what might occur; rather each universe state is formed for the first time without prior indication only at the moment it comes into being. If we were to look at any moment in a pure indeterministic universe with our ontic magnifying glass we would find nothing there to indicate what the future might bring. We find nothing in the fabric of the relata that tells us, or even hints at, what the next moment will contain and how it will be shaped.

[ONE PATH] A universe, as it evolves, takes only one actual path from moment to moment.

We can now begin to unpack our issue. We ask what ontic features of a universe, present at some given moment, indicate that it is offering alternative paths in the future. It would seem that **[DET]** suggests that determinism does not offer alternative paths and so cannot provide an answer.[1] And **[ONE PATH]** says that, as the universe only takes one path, alternative possibilities that might

[1] Part of what is to come in this chapter stands as a more careful arguments for this claim.

have been mooted in discussions are not realised and so have no ontic presence at any point following the realisation of the actual path.[2] And further **[INDET]** suggests that these unrealised future paths are not present in the fabric of the indeterministic universe at any time prior to or at the moment in question.

What this suggests is that our search for an ontic construal of choices, as possible alternative futures, is doomed to failure. It would seem that neither determinism nor indeterminism can provide us with a resolution. It is the purpose of this chapter, then, to argue more carefully that this conclusion is correct and that we must introduce the notion of immanence into our models to provide the ontic underpinning for the choices we need to construct our version of the exercise of free will.

Before we recast the problem of ontic choice in terms of counterfactuals, we shall make some brief informal comments on the seeming clash between what we have said here and the usual notion of choice as a form of indeterminism. We explain that what is usually taken as an understanding of pure indeterminism is in fact one that smuggles in the notion of immanence.

We do this by way of a simple example. Imagine a billiard ball moving along in an indeterministic universe. The usual version of indeterminism would be to say that the billiard ball might evolve down one of a [finite] number of possibilities. Perhaps it will continue in a straight line, or perhaps it will turn left, or perhaps it will turn right. If we claim that these are the only three possibilities for the billiard ball but also claim that which it will perform is not set in advance, then it would seem that we have a valid version of indeterminism. In fact we do indeed have a valid version, but what is less clear on first glance is that we have smuggled into our story a little more than we have placed into our models so far. We have made the assumption both that the billiard ball will continue to exist and that it has only three possible paths open to it. However, if that is the case then we must find these possibilities written in the fabric of the universe prior to one of them actually occurring: there must be some ontic indication in the present that these are the only possible future scenarios. But as we mentioned above in **[INDET]**, there is no such ontic presence in purely indeterministic universes so we must have smuggled something extra into our model. As we shall explain at the end of the chapter, what we have smuggled in is an assumption of immanence. We have assumed that there are only a finite set of possible futures and that these

[2]We ignore a complication here: it might be argued that if we allow the universe to spilt or branch into [causally distinct] pieces and our conscious awareness - and bodies - to split into a set of beings, one strand in each piece of the split universe, then each strand of consciousness would experience a future different from that experienced by other strands. We ignore such many-worlds type scenarios - see, for instance, Saunders et al. (2012) and Bohm & Hiley (1993, Ch. 13) - because they do not provide alternative futures but rather a single future under our definitions. It might be the case that there are ways of making such many-worlds scenarios sufficiently sophisticated with respect to definitions of what constitutes a self, or selves, to get some purchase on choice but it is not clear how this can be achieved in a manner that would allow us to satisfy the two criteria for the operation of free will. We do not pursue many-worlds in this study.

are all somehow ontically lurking in embryonic form in the present moment. They are embryonic because eventually one will become actual, as it is realised in the future, and the others will fizzle away unrealised into the universe. The point is that there is an understanding of indeterminism that does have immanence built into it: this immanence goes unnoticed in the usual analyses.

4.3 The problem and our strategy: introducing the counterfactuals [X] and [Y]

Our problem can be set out formally as follows:

We are given that [relatum] A actually occurs at time t_1 and causes B to occur at t_2.[3] We then want to know what ontic features present in the universe at or just prior to t_1 make it true **just before or at time** t_1, where 'just before' means 'at t_0', that something else could occur in place of this actual occurrence of A-causing-B. In other words we want to understand what ontic factors **extant just before or at time** t_1 make at least one of the following time-bound scenarios a genuine possibility, and not merely a conceptual possibility:

[X] P occurs at time t_1 in place of A and causes Q at t_2 in place of B

[Y] A occurs at time t_1 but causes C to occur instead of B.

We assume, for ease, that P is not entailed by A or vice versa.[4] We shall call these 'counterfactuals' even though we are considering them at time t prior to their becoming counter to facts.

We shall proceed with our analysis of **[X]** and **[Y]** in three stages. First, we look at the standard approach to dealing with counterfactuals using Lewis's possible worlds,[5] as outlined primarily in Lewis (2001a).[6] Lewis's account is for Humean universes,[7] or at least those worlds governed by laws which are systematizations of patterns rather than ontic laws. We, in contrast, are looking at connexion-bound universes, but we start with Lewis's approach as it is a standard one for dealing with counterfactual and it provides motivation for our subsequent discussions. We put the case that this approach is not adequate for our needs for a number of reasons: primarily because it requires that to answer our question, posed in counterfactual terms, using Lewis's methods we must refer to non-actual universes. We have set ourselves the task of seeking what it is in our [actual] universe considered alone that gives ontic leverage to choice so referral to possible non-actual worlds gives

[3]We do not continue to refer to relata as 'relatum P' etc; rather we just write 'P' etc.

[4]It should be clear that we also assume that all relata are positive in the sense that we do not countenance 'not A' as a valid relatum.

[5]We do not consider alternative approaches, such as two-dimensional semantics, here. See, for instance, Chalmers (2006).

[6]There is an interesting related discussion in Stalnaker (1976). It is also worth consulting Lewis subsequent publication and responses to criticisms, many of which can be found in Lewis (1987).

[7]For a discussion of laws and Humean universes, see Urbach (1988).

us no ontic leverage. So we reject Lewis partly because he must refer to universes that are causally disconnected from our actual universe, even though they may well be real in some other sense. Recall here that we defined, in section 3.9, a universe as a closed causal nexus. Any other possible world that is causally isolated from our [actual] world has no ontic purchase on the machinations of our universe - recalling causality is the only diachronic ontic relationship between relata

Further, if we were to attempt to move Lewis's methods to apply within a single universe then we should have to pay careful attention to the ontic structure of the single universe. And we shall observe that if we take the methods Lewis outlines across to apply within a single universe then they stand in tension with this restriction: they are somewhat *ad hoc* as to which [possible] universe we must take as our marker for the truth of some counterfactual but we cannot assume such an *ad hoc* approach reflects the restrictions that arise from staying true to the ontic nature of the actual universe.[8] So Lewis's ideas of similarity across possible worlds taken as a guide to permissible modification within a universe are not precise enough and might not respect the ontic features of the actual universe. And, if we are to confine our analysis and our results to the actual universe alone, permissible changes must be measured in terms of what we have in front of us alone. That is, we have no beef and make no comment upon the plausibility of Lewis's scheme as a conceptual analysis of counterfactuals but have less faith in it as a potential guide to an ontic analysis.[9]

We then move on to investigate this second issue, that of what we might legitimately change within a universe when exploring our issue in counterfactual terms, by introducing the idea of **causal permissibility**. In simple terms this rallies the ideas we have outlined in previous chapters to investigate what we might be able to alter in our actual situation, namely the actual scenarios where A causes B, whilst staying faithful to all the ontic baggage that comes along with a universe and a scenario. In doing so we show that considerations about scenarios of type **[X]** reduce to the considerations of scenarios of type **[Y]**. However, whilst this simplifies what aspects we must investigate, it does not remedy the issue as to what occurrences in the actual universe make **[Y]** type statements true possibilities at some time. We go on to argue that **[Y]** type statements cannot help to resolve our problem, regardless of the type of universe that they inhabit. That is we conclude that counterfactuals cannot be countenanced as genuine possibilities if we stay true to the ontic features of the structures we have set out so far.

[8] Lewis (1979) does set out some general principles for this but they are not precise. He says, on page 472, '(1) It is of the first importance to avoid big, widespread, diverse violations of law. (2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails. (3) It is of the third importance to avoid even small, localized, simple violations of law. (4) It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.'

[9] We hold Lewis's methods can only amount to a conceptual analysis, and not an ontic one, as they depend on an epistemic construal of laws and not an ontic construal.

We then move to argue that **[Y]** can only have ontic underpinning in the actual universe when the relata that **[Y]** concerns are of a very specific sort. They must be immanent relata, something we define formally in section 4.6. Immanence will be what gives us a clear understanding of the claim a universe presents us with genuine choices and it will then form a central part of the model we propose in Chapter 5 for the operation of free will.

4.4 Lewis's 'Possible Worlds'

Here we start our steps towards a solution to our problem with a brief look at Lewis's possible world semantics. We do this primarily to motivate the next step we take, that of investigating how ontic constraints might affect the moves we make within a universe when working towards an ontic notion of choice.

As we hinted above, Lewis's approach cannot possibly be an answer to our problem as he seeks to eke out the conceptual baggage that accompanies counterfactuals - recall, we take his analysis as conceptual as he bases it on an epistemic construal of the laws of nature. Indeed, Lewis (2001a) takes his views of laws from Ramsey,[10] and he states,

> I adopt as a working hypothesis a theory of lawhood held by F.P. Ramsey in 1928: that laws are "consequences of those propositions which we should take as axioms if we knew everything and organised it as simply as possible in a deductive system." (Lewis, 2001a, p. 73)

This is an epistemic approach to laws, as explained in section 2.10.1, and is bound-up with the assumption that the world is regular - it behaves itself enough to allow our finite theories to have some leverage. Recall, in section 2.10.1, we explained that an epistemic approach to laws is one where the regularities observed in the world are summarised in finite, usually mathematically-couched, theories. We also noted there that we cannot claim that the theories reflect actual ontic links - the web of connexions - between relata.

Lewis sets up a system to interpret counterfactual statements of the form 'if P had occurred then Q would have occurred' and writes this in symbolic from as 'P $\mathrel{\Box\!\!\rightarrow}$ Q'. We shall simplify some technical aspects of Lewis's approach to make the discussion less involved.[11]

The way to interpret these counterfactuals, according to Lewis, is to consider the embedding of the actual world in a set of possible worlds and then to examine how the counterfactuals pan out

[10] The Ramsey quote, which appears on page 73 of Lewis (2001a), can be found on page 42 of Ramsey (2013).

[11] For instance, we do not look at his limit assumption: Lewis (2001a, p. 19 ff).

in this augmented realm.[12] That is we start with a structure, perhaps like that used in $S5$ modal logic,[13] with the possible worlds that envelop the actual world - the accessible possible worlds - satisfying some explicit criteria. Lewis says,

> We call these worlds *accessible*, meaning thereby simply that they satisfy the restrictions associated with the sorts of necessity under consideration. Necessity is truth at all accessible worlds, and different sorts of necessity correspond to different accessibility restrictions. (Lewis, 2001a, p. 5)

Once we have the set of possible worlds, we order them around the actual world. We order them by making worlds that are more similar to the actual world closer than those that are less similar - with the idea of similarity left a little vague initially. We then mark all the worlds where P is followed by Q according to law, that is worlds where both P and Q occur in place of A and B and P and Q are linked by laws. We also mark worlds where P occurs but Q does not occur. Amongst these two sets of marked worlds we can seek the one that is most similar to our actual world,[14] and if this is one where both P and Q occur at appropriate spatio-temporal locations, then we deem the counterfactual to be true; and if it is a world where P occurs and Q does not, then we deem the counterfactual to be false.

For instance, to take Lewis's initial example concerning kangaroos:

> 'If kangaroos had no tails, they would topple over' seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over. (Lewis, 2001a, p. 1)

If it is the case that when we consider all the worlds that are accessible from our world - perhaps with accessibility defined as 'obeying the same laws of nature' - that the worlds where kangaroos have no tails and topple over are more similar to our world than those where they have no tails and stay upright, then the counterfactual is true. Of course, this sort of counterfactual is not like the ones we are concerned with. We are looking only at specific events at specific times. We are not concerned with talking of generalities, of kangaroos or cabbages and kings on Tuesdays and Thursdays and the like.

---

[12] The term 'world' is a synonym for 'universe' here.

[13] $S5$ modal logic and related issues are discussed in both Forbes (1986) and Cresswell & Hughes (1996).

[14] We shall, for simplification assume that this is well-defined. If not, Lewis has further methods using sequences of worlds that amount to the same idea. We also assume, for simplicity, that P is not made impossible by the accessibility criteria.

The similarity criterion is complicated and somewhat *ad hoc*. It is a trade off between keeping the past and surroundings fairly unchanged and keeping the laws of nature fairly unchanged. Lewis says,

> We might think it best to confine our attention to worlds where kangaroos have no tails and *everything* else is as it actually is; but there are no such worlds. Are we to suppose that kangaroos have no tails but that their tracks in the sand are as they actually are? Then we shall have to suppose that these tracks are produced in a way quite different from the actual way. Are we to suppose that kangaroos have no tails but that their genetic makeup is as it actually is? Then we shall have to suppose that genes control growth in a way quite different from the actual way (or else that there is something, unlike anything there actually is, that removes the tails). And so it goes; respects of similarity and difference trade off. If we try too hard for exact similarity to the actual world in one respect, we will get excessive differences in some other respect. (Lewis, 2001a, p. 9)

He accepts that if kangaroos were to have no tails then we must rule out worlds with identical past histories to ours: tailless kangaroos will have a different genetic makeup, leave no tracks in the sand and so forth. That is, his similarity does not require identical past histories. This may not be useful for our single-world ontic cases as we are seeking something in the actual world that makes time-bound counterfactuals true and changing the past dramatically would take us too far away from our actual world. This is not a criticism of Lewis if we take him as seeking a conceptual understanding of counterfactuals and not an ontic one but for us it marks a point of difficulty.
Another example that Lewis considers is counterfactuals in deterministic universes. Here, intuitively at least, it might seem odd to allow laws to alter as we move from one world to another. But, as Lewis explains, adjusting a few laws locally is much more palatable as a criterion of similarity than restructuring the whole past. He says,

> Suppose a certain roulette wheel in this deterministic world *i* stops on black at time *t*, and consider the counterfactual antecedent that it stopped on red. What sort of antecedent-worlds are closest to *i*? On the one hand, we have antecedent-worlds where the deterministic laws of *i* hold without exception, but where the wheel is determined to stop on red by particular facts different from those of *i*. Since the laws are deterministic, the particular facts must be different at all times before *t*, no matter how far back . . . On the other hand, we have antecedent-worlds that are exactly like *i* until *t* or shortly before; where the laws of *i* hold almost without exception; but where

> a small, localized, inconspicuous miracle at t or just before permits the wheel to stop on red in violation of the laws. Laws are very important, but great masses of particular fact count for something too; and localized violation is not the most serious sort of difference of law. (Lewis, 2001a, p. 75)

And goes on to say,

> . . . some of the antecedent-worlds where the law is violated may be closer to i than any of the ones where the particular facts are different at all times before t. (Lewis, 2001a, p. 75)

So for Lewis, preserving the great mass of the past is a heavier concern than worrying about small local violations in laws when deciding which worlds are to be close and so which are then to be the underpinnings of the counterfactuals.

Lewis's scheme furnishes a plausible candidate for analysing the conceptual basis of counterfactuals. And recall, we decided to look briefly at Lewis to see where we might go with our ontic analysis. However, recall also that we seek some ontic features of the *actual* world that provide the presence of possibilities prior to their actualisation or disappearance. We must seek our solution *within* our world and we have no remit to wander outwards peering into alien universes. We cannot even find succour in taking other possible worlds as being real but elsewhere. If they are elsewhere they cannot be here, they cannot have any causal influence over what happens in our actual world - a world, recall, that is causally closed.[15] So we take from Lewis the idea that we might be able to look for a solution to our problem by looking to counterfactuals but doing so whilst keeping the past fixed and changing parts of the structures in our world, but changing them only locally. And additionally, we must stay firmly within our own universe and deal with its ontic structures alone. It is to do this we now turn. We seek to investigate if we can find a solution to our problem by investigating how we might understand our counterfactuals within our actual universe with all its accompanying ontic baggage.

4.5 Causal permissibility

In the next few sections, 4.5.1 to 4.5.8, we set out to investigate the issue we outlined above, namely what we might change within a universe whilst remaining true to its ontic structure, by paying heed to the causal structures from Chapter 3. This will provide us with our intermediate step as we move away from possible worlds and look inwards to a single universe for an ontic

[15]We deal with the special case of separate worlds ontically supervening in Chapter 6.

underpinning to choices.

We shall take this intermediate step by outlining a theory of causal permissibility:

> **DEFINITION 4.1: Causal permissibility**
>
> We say that some proposed change or alteration in relata or connexions at some moment in a universe is **causally permissible** if it is consistent with the ontic structure of the universe under consideration.

Here we set out what this consistency is to entail.

In broader terms, causal permissibility is an attempt to encapsulate what is allowed and what is forbidden by the ontic structure of a universe when seeking genuine alternative future scenarios. It looks inwards at the restrictive structure of the world to see what this allows when considering how we might understand the possibility of **[X]** and **[Y]**.

4.5.1 Arguing [X] reduces to [Y]

Recall **[X]** and **[Y]**:

> **[X]** P occurs at time t_1 in place of A and causes Q in place of B
>
> **[Y]** A occurs at time t_1 but causes C to occur instead of B.

In this section we argue that **[X]** reduces to a form of **[Y]** if we hold on to some simple assumptions. Initially we shall assume we are in a universe that does not have any causal over-determination and we dismiss the relevance of spontaneous events or epiphenomena. We deal with over-determination in section 4.5.5 below.

First we shall consider a simplified version of the actual occurrence that relatum A, at time t_1 causes relatum B at time t_2. We illustrate this in Figure 4.1. The nodes represent relata, possibly fundamental, and the arrows represent general connexion-bound causal relationships between the relata. We note here that it is possible that A might, in addition to being the complete cause of B, also contribute causally to other relata.

Our counterfactual **[X]** requires us to replace A by P and B by Q as in Figure 4.2. We have replaced $\mathcal{S}_7(t_2)$ and $\mathcal{S}_9(t_2)$ as they may well be different from the original relata.

We can use these diagrams to understand the basic issues we face in constructing counterfactuals. We start by asking three questions about replacing A by P:

> **[q1]** Can we legitimately replace A by P?
>
> **[q2]** Does P cause Q?
>
> **[q3]** How disruptive to other relata at t_2 is P?

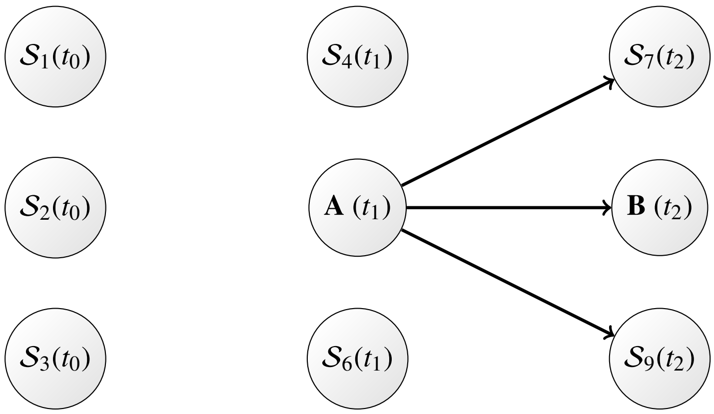

Figure 4.1: A actually causes B

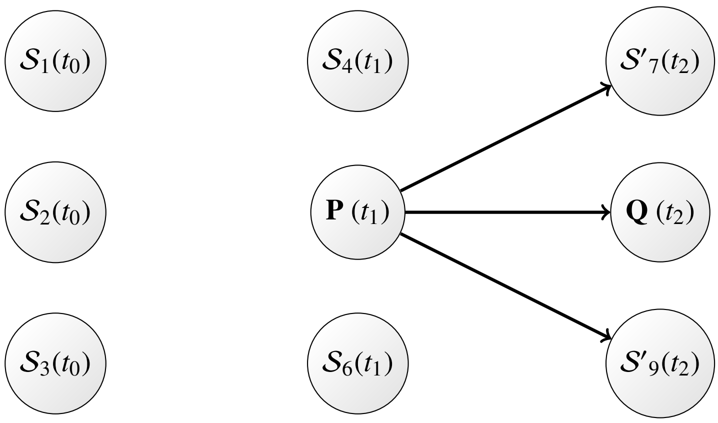

Figure 4.2: Counterfactual: P causes Q

We consider the future changes that occur by replacing A by P to be unimportant. This is because we are interested in replacing A by P and then seeing what happens. We expect that the replacement will then result in Q's happening and that is our sole concern. We do not need to concern ourselves with other future changes that may arise from our adjustments: we ignore the possibility that the there might be other dramatic changes elsewhere in the future. That is, we need not worry about the possible changes in relata at t_2. We do not need to concern ourselves with **[q3]**.

The question as to whether P causes Q is more difficult to resolve. In section 2.10 we discussed the various forms that ontic laws can take. There we explained that there is no restriction on the way the laws operate from moment to moment and place to place. That is, it is quite acceptable to have ontic laws that necessitate P causing Q at t_c but not at t_b or t_k etc. If we are given the relatum P we cannot say with certainty that it will cause Q unless there are further restrictions on both the relata. We need to be able to identify relata of type P and Q and the ontic laws so that we can guarantee they necessitate the causal relationship between P and Q.[16] To launch the possibility of general counterfactuals we must impose two restrictions which derive not from the relationship between A and B but from the purported counterfactual relationship between P and Q. We might propose:

> **[RES1]** Relata divide into types that are identifiable;
>
> **[RES2]** Ontic laws must necessitate spatially and temporally constant relationships between relata.

However, this is a little too restrictive as **[RES1]** and **[RES2]** when applicable to all laws and relata require the universe to be deterministic. This is because if all relata can be categorised into types, by **[RES1]**, and the ways they cause are set by the type that they are, by **[RES2]**, then they will ontically necessitate their own causal consequences - the condition for a universe to be deterministic. It would, however, be sufficient to restrict counterfactual statements to those that involve relata and concomitant laws that satisfy **[RES1]** and **[RES2]** but not require all relata and all laws to obey these conditions. All we need for our arguments here is for the restrictions to be valid when counterfactuals are considered valid.

So, we require:

> **[RES1a]** Some relata divide into types that are identifiable;
>
> **[RES2a]** Ontic laws applicable to these relata must necessitate spatially and temporally constant relationships between the relata.

[16] We note that this restriction is not required if we take P and Q as the negations of A and B but we do not consider that B's not occurring is a valid relatum specification here.

We shall assume this is the case in what follows. It may be that **[RES1a]** is really a consequence of our need to *articulate* alternatives rather than one that is ontically necessary. This distinction is not important for what follows, so we shall just assume that **[RES1a]** and **[RES1b]** are what is needed to allow us to claim P causes Q and worry no further about the relationship between P and Q.

We now turn to discuss **[q1]**. First, we put an absolute restriction on counterfactuals: we require that in replacing an actual occurrence A with a counterfactual occurrence P we keep both the relata prior to A and those contemporaneous with A unchanged. That is we do not seek to change the past or the present in any manner in our counterfactual beyond the replacement of one relatum with another.

For instance, take the classic time-bound counterfactual that says, "if I had lighted the match at time t then the forest would have burnt down at t+1." In saying this we do not assume that my history, or that of the surroundings, was to be changed in any way prior to or at time t. We merely seek to say that a different action at time t, with an identical past to the one that actually occurred, would have given different subsequent occurrences. As we are setting our arguments within a single universe we could sanction that events in the present can reform that past in some manner. If we allowed such scenarios we would have to work with models that allowed for a form of backwards causality. Here we have rejected this possibility and so we are justified in making our assertion into a formal requirement of counterfactuals:

> **[RES3]** Counterfactual antecedents must be compatible with the same past and the other present relata and connexions that accompany the actual relata which the antecedent replaces or those subject to **[RES4]** below.

We can see this in the transition from Figure 4.1 to Figure 4.2 where the past relatum, at t_0 and other current relata surrounding A at time t_1 appear in Figure 4.2 surrounding P.

Now we shall argue that, if we accept that P can indeed cause Q - so that Q is perhaps an essential consequent of the occurrence of P - then **[X]** must reduce to a counterfactual of type **[Y]**.

The argument is simple. If we allow our assumption that we can hold P causes Q then we must ask if the replacement of A by P is legitimate. In other words, assuming here that $\mathcal{S}_2(t_0)$ causes A, we must ask if we can maintain $\mathcal{S}_2(t_0)$ whilst changing A to P. This is clearly a version of **[Y]**.

In the next section we put forwards what seems to be a plausible argument that this can only be achieved in circumstances where the connexion between $\mathcal{S}_2(t_0)$ and A is separable; that is, if $\mathcal{S}_2(t_0)$ is a separable relatum. We put this argument as it both seems plausible and as it appears, in one guise or the other, in everyday talk. Once we have outlined the argument we shall go on to explain,

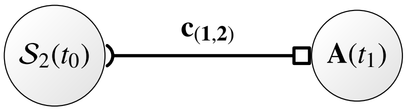

Figure 4.3: Necessary essential relatum and connexion

in section 4.5.6, why it is flawed as an argument. There we shall argue that separability is not a route to allowing individual relata to be modified or replaced and we shall use arguments that echo those we advanced in section 3.9, but here applied to individual relata rather than a universe as a whole.

4.5.2 The causal permissibility of [Y]

In this section we deal solely with connexions that are necessary - locally deterministic - and we postpone dealing with indeterministic scenarios until later.
As mentioned, the argument we shall advance here, whilst plausible, is partly incorrect. It will act as a bridge to our discussion in section 4.5.6; there we revisit the idea of separability, first introduced in section 3.9, and show it is far more restrictive than we assume in this section. Recall, in section 3.9 we said of relata separability:

> Informally, this says that to be separable a relatum must ontically necessitate its future relata but the manner in which it does so, and hence the way it impinges causally on the future, can vary and so is independent of the nature of the relatum in itself - that is, independent of the relatum's causal content.

And in this section we take this to give licence to making local changes without wider repercussions: that is, when faced with a separable relatum we take it that we can change what it causes in the future, by changing the corresponding connexion, without effecting changes elsewhere. It is this assumption that will eventually be shown to be in error when we come to revisit, in section 3.9, the arguments we advance in this section.

We start by considering the case where the connexion between $\mathcal{S}_2(t_0)$ and A is both necessary and non-separable, as illustrated in Figure 4.3. Here we are not in a position to change the nature of the ontic connexion between $\mathcal{S}_2(t_0)$ and A. That is to say that A cannot be replaced without disputing $\mathcal{S}_2(t_0)$ and this violates **[RES3]**. As a result, in cases where the connexion to A is essential, no counterfactual can be countenanced. This is a correct result and need concern us no further.

Next consider a case where A is caused by $\mathcal{S}_2(t_0)$ and the connexion is separable as in Figure 4.4.

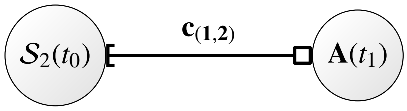

Figure 4.4: Necessary separable relatum and connexion

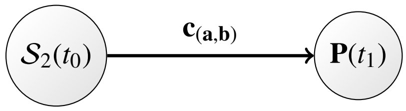

Figure 4.5: General causal connexion to give P

If P is a permissible replacement for A then we must argue that there is some notion of causality, denoted by a plain arrow in Figure 4.5 that is a legitimate replacement for the causal relationship in Figure 4.4.

We require that altering a connexion in this way is restricted:

> **[RES4]** Connexions may be altered in counterfactual situations if their alteration has no implications on the nature of relata or connexions contemporaneous to or prior to those that are altered.

We must be careful in using **[RES4]** as there are two ways we can interpret what it requires. It may be the case that whilst connexions are separable in a universe, the same connexion type occurs after every relatum of a given type so that the relationship between relata-type and connexion is uniform throughout the universe. In this case we might take it that changing a single separable connexion in one place is a violation of **[RES4]**. However, we shall not take this line because we have not taken separability of relata to be a doctrine that is relevant universally; rather, recall our notion of separability for relata is one that applies to each relatum individually. We take **[RES4]** to say that separability must be achieved by altering a single connexion alone and have no repercussions on other contemporaneous or prior relata or connexions.

Now if we take separability to be merely the assertion that a connexion can be considered separately from its corresponding antecedent relatum, then there seems to be no reason not to allow the connexion to be changed in isolation. That is, we seem to be in a position to conclude that if $\mathbf{c}_{(\mathbf{1},\mathbf{2})}$ in Figure 4.4 is the connexion of a separable relatum then it is causally permissible to change it to $\mathbf{c}_{(\mathbf{a},\mathbf{b})}$ of Figure 4.5 and hence effect the change we require for **[Y]** to be causally permissible.

We might think of this as something akin to Lewis's minor miracles that we met above. However, there is a slight difference as Lewis's miracle was a deviation from the laws that applied universally and here we merely allowed a separable connexion to be altered in isolation in line with its nature as separable. We have justified this as we have taken ontic laws and epistemic laws to be different. Here, it may well be the case that altering a single separable connexion violates the patterns that are evident in the universe, and hence stands as as an exception to prior epistemic law. But, in contrast, the alteration is itself a local - i.e. at a single relatum and at a specific time - re-framing of ontic laws and that here is allowed according to our interpretation of separability and our arguments. And it is precisely what causal permissibility seeks to encapsulate.

So, we have a result that obeys **[RES4]**: that A may be replaced when its cause connects to it by a separable necessary connexion.

Below, in section 4.5.6, we shall argue that the interim conclusion we have reached here is incorrect. We argue there that our scenario here - namely holding to the causal permissibility of replacing connexions stemming from separable relata - is incorrect: **[Y]** is not causally permissible when ontic necessity is involved in connecting an antecedent to a consequent, regardless of whether the connexion is separable or not.

Whilst, in this case, we have 'established' that a certain connexion type permits A to be replaced, we have not investigated what relata are permissible as replacements. It does seem that, as the connexion we are replacing is isolated, then any connexion is permissible and any consequent P is permissible. We have already argued above that P must fall under a relatum type - perhaps for purposes of articulation - but it does not follow that it must be of a type in this scenario. We shall not comment on the restriction on P as it is usually specified as part of the counterfactual antecedent. If it is the case that our specification is forbidden by some other features of the relata $\mathcal{S}_2(t_0)$ or the ontic connexion that emanate from it then our counterfactual fails to launch and questions of its truth become irrelevant.

Before moving to explain why the arguments of this section cannot stand, we take a minor diversion in the next three sections to clear up a few, slightly peripheral, technical matters.

4.5.3 A brief aside on hybrid causes

In this section, we comment briefly on hybrid situations concerning **[Y]**. We do this for completeness.

Consider a case where A is caused by a set of relata via a varied set of connexions. We illustrate one such case in Figure 4.6. The arrow from $\mathcal{S}_1(t_0)$ to $\mathbf{A}(t_1)$ with a rhomboid head is used to denote an indeterministic causal connexion.

Here, in accordance with our discussions above, we must keep the exact connexion between $\mathcal{S}_3(t_0)$

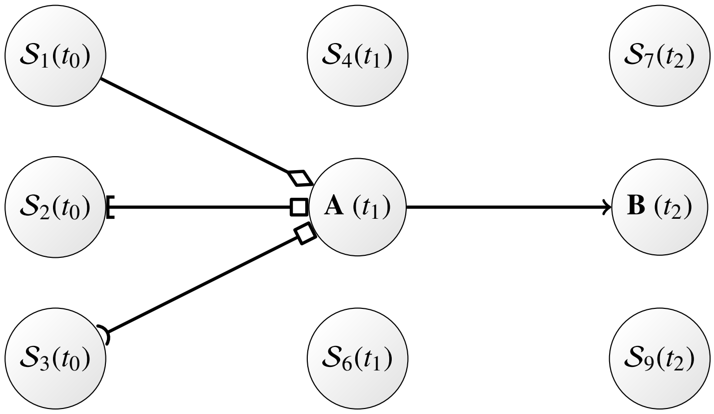

Figure 4.6: A is caused via a variety of causal connexions

and A, we may vary that between $\mathcal{S}_2(t_0)$ and A so long as the connexion is unique and not bound by being a member of some type, and we have free reign over the connexion between $\mathcal{S}_1(t_0)$ and A. So what we may replace A with, namely P, is partly restricted by this set of connexion types. So whilst the arguments we have put above stand in this case, P must also respect the causal aspects that derive from $\mathcal{S}_1(t_0)$ and $\mathcal{S}_3(t_0)$. This might mean that adjustments in the connexion from $\mathcal{S}_2(t_0)$ are less powerful in what they might achieve.

It is certainly possible to continue with more and more complex hybrid situations but as they will play little part in what follows, we do not pursue them here.

4.5.4 Restrictions from measure preservation

In this section, we comment on some additional restrictions that come from measure preservation - restrictions we shall just assume are fulfilled.

If we return to examine the replacement of A by P as illustrated in Figures 4.4 and 4.5, we must be aware of a further restriction on P. If it is to replace A and if we are considering situations where measure is preserved by connexions then the measure of the causal content of P must be equal to the measure of the causal content of A, and further the measured amount of contribution from $\mathcal{S}_2(t_0)$ must be the same in both cases. That is we require:

$$\mu(P(t_1)) = \mu(A(t_1)) \tag{4.1}$$

and also

$$\mu(\mathcal{S}_2(t_0), \mathbf{A}(t_1)) = \mu(\mathcal{S}_2(t_0), \mathbf{P}(t_1)) \tag{4.2}$$

Equation (4.1) requires that the amount of ontic content of A and P are the same so that replacing A by P does not violate measure preservation - there is no loss in the total quantity of causal content in the universe by replacing A by P. Equation (4.2) follows from (4.1) when measure is preserved. It says the amount of contribution from $\mathcal{S}_2(t_0)$ must be the same in both the case that it causes A and in the case when it causes P. As such (4.2) is a restriction on the connexion between the states.

For instance, were we to consider the counterfactual, 'if the sun were to become twice as massive at time t, then the earth would be pulled out of orbit' we would count it as invalid if it were the case that μ measured energy content. This is because replacing the sun with one twice its mass violates (4.1) and hence (4.2). Of course, the everyday use of this counterfactual makes sense but that is because it can be thought of as a general comment on the epistemic laws of nature rather than a time specific claim of change. It is really saying that a sun twice as massive as ours would not allow earth to orbit in its current manner according to current [epistemic] law. If, however, we consider it as we have, as a specific time-bound statement applicable as a counterfactual to a specific actual occurrence it might be invalid.

4.5.5 Dealing with over-determination

Our focus here is on understanding the legitimacy of replacing A by P under various ontic scenarios. Here we consider two cases: one where A is fully causally over-determining and one where it is fully causally over-determined. We do not look at hybrid cases where there is some partial over-determination.

First consider a case where A causally over-determines some future relatum. We illustrate this in Figure 4.7 using dotted lines to indicate over-determination. We do not make an issue of the connexion type here. We have B caused by both A and $\mathcal{S}_6(t_1)$.

Now if we replace A by P then B still occurs, as it is caused by $\mathcal{S}_6$ which is left untouched at t_1 but now P also will cause Q at t_1 and contribute to other relata too perhaps, including B. This may be awkward but it is not forbidden by any of the permissibility rules we have proposed and so it would seem not to be an issue of concern. It might, however, violate measure preservation but this is not a firm condition for permissibility.

For instance, consider the example of two assassins that both shoot dead an evil pirate.[17] Now let us consider the counterfactual situation where we replace one assassin's bullet-firing gun with

[17]Where we assume there is no problem with this an example of causal over-determination. See our discussion in section 3.6.1.

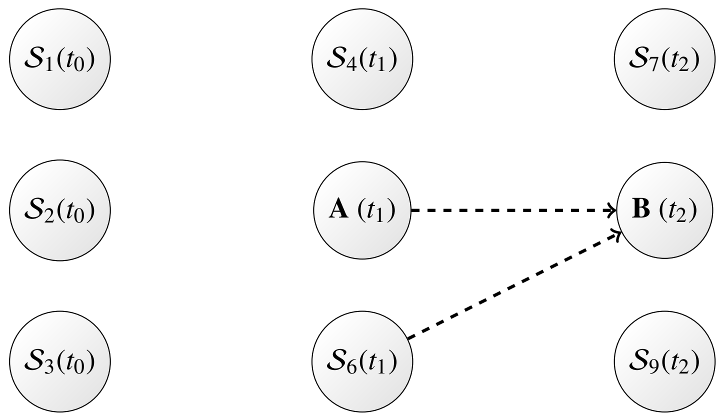

Figure 4.7: A causally over-determines B

a paint gun. In that case the evil pirate is still killed but he is also covered in paint. So the counterfactual we would propose in this case maintains the B-event, namely the pirate's death, but adds to it in some manner - namely by splattering the pirate with paint. So here we might countenance counterfactuals of the form, ‘if one assassin had used a paint gun at time t the evil pirate would still have died at t+1’ or, ‘if one assassin had used a paint gun at time t then the pirate's corpse would have been paint splattered’. The point is, there is no great difficulty in dealing with the antecedent of a counterfactual being causally over-determining. In such cases, the issues remain as discussed above and concern the connexions that lead to A, not from A.

Now, consider a second case where A is itself causally over-determined, such as illustrated in Figure 4.8. Here, replacing A puts the same restrictions on the over-determining connexions as above in the hybrid cases. For instance, if the connexion from $\mathcal{S}_2(t_0)$ to A is essential [i.e. $\mathcal{S}_2(t_0)$ is an essential relatum] whilst that from $\mathcal{S}_3(t_0)$ is indeterministic then P cannot replace A but must rather be a relatum that entails the full content of A caused by $\mathcal{S}_2(t_0)$ as well as some additional content permitted by adjustments in the connexion from $\mathcal{S}_3(t_0)$. It is not easy to speculate what restrictions this may place on the possible P replacements for A. We leave it as an open question.

Over-determination is not an issue for permissibility in counterfactuals when the actual occurrence being replaced is over-determining, unless measure conservation is also required, but it does place complications and restrictions on the counterfactual antecedent when the actual occurrence being replaced is over-determined. As we shall not be concerned with causal over-determination in counterfactuals we leave the matter here.

4.5.6 Dismissing local separability in causal permissibility

In this section we return to the arguments we advanced in section 4.5.2. Recall, there we argued that when faced with separable relata we can adjust their causal consequences with impunity. That

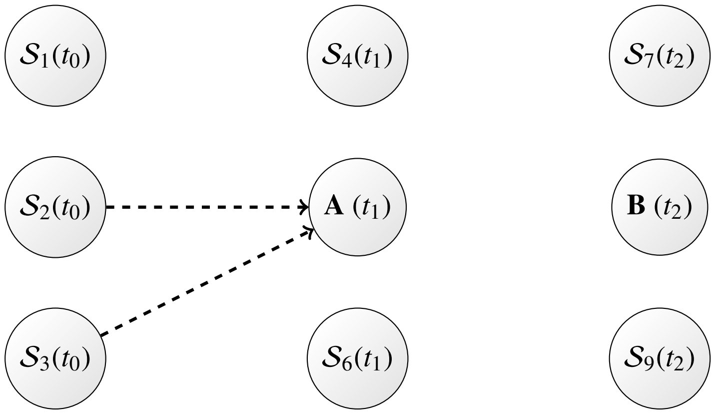

Figure 4.8: A is causally over-determined

is, we can change the universe locally around a separable relatum by modifying a connexion and we can do so in a manner that does not have wider repercussions.

Recall we said there of our definition of relata separability that,

> ... we take this to give licence to making local changes without wider repercussions: that is, when faced with a separable relatum we take it that we can change what it causes in the future, by changing the corresponding connexion, without effecting changes elsewhere.

Here we shall argue that this encapsulation is not correct. Indeed, separating relata from their connexions is not an innocuous move and the sentiment expressed in the quote from section 4.5.2 is in error.

Our discussion here also extends the discussion we started in section 3.9 to encompass separable relata in a universe that is not completely deterministic. We shall show that separable relata, in either a deterministic or indeterministic universe, cannot be separated from their causal consequences and their connexions without wider repercussions. What this amounts to is the claim that separability, whilst giving conditions on individual relata, is not a local feature of individual relata - local in the sense that it affects only the relatum under consideration - but rather a property that is inherited from global properties of the universe. This means that we cannot countenance minor local miracles, as violations of ontic laws, *within* any sort of universe: minor miracles are not causally permissible.

Our argument is advanced in stages: first we set out an example, drawing on the causal structures of Chapter 3, that shows how separability is often taken to operate as we cross to possible worlds. This example then helps to ground some of the subsequent discussion. We then put the case that we cannot blithely separate relata from their connexions locally in deterministic universes and so conclude **[Y]** type counterfactuals are not causally permissible in such scenarios. We then move to show that our arguments apply to separable relata - which, recall, always ontically necessitate their own causal consequent - in hybrid universes. A universe is hybrid if it contains both deterministic and indeterministic aspects. And so we must also conclude that **[Y]** type counterfactuals are not causally permissible in hybrid universes.

First, then, an example: it is well known that an electron moving at right angles to a uniform magnetic field will move in a circle, or circular arc.[18] In very simple terms, if we consider the corresponding relata to be separable then we can imagine transiting to a possible world, keeping both the magnetic field and the electron, whilst having the electron move in some other path in the magnetic field there - say a straight line or some other such arc.[19] We might then *think* of this scenario in a number of ways. We might look at the electron's motion as being the result of the combination of two causal relata - a magnetic field and a moving electron; or we might look at it as the result of a single relatum evolving, where the single relatum may be *described* as some combination of magnetic field and electron-in-motion. In moving across to our possible non-actual world, we allow the relata in some form to transit with us - so we admit the presence of an electron and a magnetic field or some such - but we think that the way they evolve and/or combine can be altered. This is the essence of separability in possible worlds.

We must be a little careful with our example, however. It suggests that we have electrons as distinct entities and can change the way they behave across the whole sweep of time as we move from universe to universe. In contrast, our concern is with individual relata and their connexions considered one at a time. What we must then do, to understand the scope of our example, is take it as applicable to a single electron considered in isolation at some moment. There is no restriction in doing this as if an electron is not an essential relatum - and, recall, we are assuming here it is separable - then the idea is that it can be altered without repercussions for other electrons at other times. What we do not consider in this work is whether there is another notion of separability where the nature of relata-types is fixed *within* a universe but can vary *between* universes. We do not consider this case as it is not relevant for the causal permissibility we are pursuing here. It is not relevant because, if this global separability were to apply, then we would be forbidden from changing the nature of a single electron at some moment in the history of the universe without changing the nature of all electrons past, present and future: such extravagance is forbidden by our rejection of backwards causality.

[18]For instance, this is dealt with in Feynman (2011) and all elementary physics texts on electromagnetism.

[19]This might be challenged. For instance, it might be argued that the relationship between the motion of the electron and the direction of the magnetic field must be mathematically simple. If that is the case then it is possible to bring in some function of the angle between the electron's velocity and the direction of the magnetic field into the epistemic laws that express the motion and construct all sorts of motions of electrons in magnetic fields.

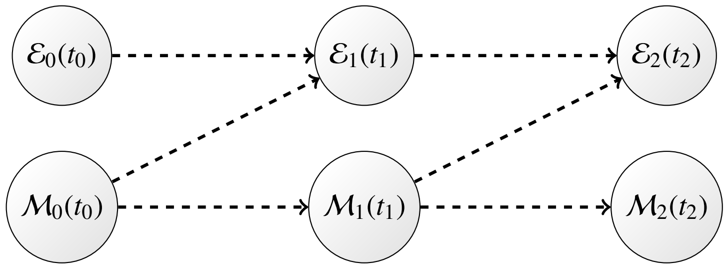

Figure 4.9: Electron and magnetic field as distinct relata in combination

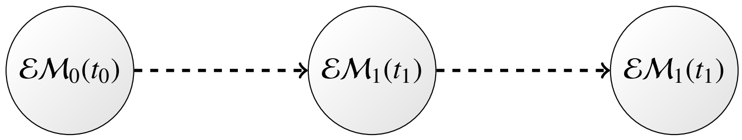

Figure 4.10: Electron and magnetic field as a single relatum

We illustrate the electron-in-magnetic-field possibilities in Figures 4.9 and 4.10.

Figure 4.9 illustrates the electron and the field considered as distinct relata combining and Figure 4.10 illustrates them as a single relatum. First, we note that the relata in both scenarios are all slightly different. For instance, in Figure 4.9 the electron relatum at t_0, $\mathcal{E}_0(t_0)$, is not the same as the electron relatum at t_1, $\mathcal{E}_1(t_1)$: they differ in their dynamic aspects as each relatum has a different instantaneous velocity. What this means is that if we are to countenance the idea of an electron [at least for our example here] then it must be the case that these relata have a common core content that persists through time, whilst allowing for additional content, such as the instantaneous velocities, to vary. If this is the case then our idea of an electron as a persisting object of some sort can be maintained even though the relata it participates in [or is built from] may vary from moment to moment. If we consider the situation in Figure 4.10 then it is less clear that we have an ontic underpinning to talk of electrons and magnetic fields - rather such terms are abstractions that we use to produce clear explanations for sets of phenomena that demonstrate suitable patterns of behaviour. As we are not concerned about the detail of which is correct we shall treat both scenarios equally in what follows.

We now turn to explain why, in a deterministic universe, separability does not allow isolated changes - and so isolated changes are not causally permissible there.

Previously, in section 2.10.3, we noted that for deterministic universes ontic laws must be carried along within the fabric of the relata so they have a presence at each moment. And in section 3.9 we explained why we rejected the idea that universes, as wholes, can be separable within our model.

For separable relata in a deterministic universe, we have three factors:

> **[Sep1]** The ontic laws are held in the current state of the universe but are set in advance;
>
> **[Sep2]** The ontic laws are not held in or governed by individual relatum content;
>
> **[Sep3]** Content-identical relata can behave, causally, in different ways

[Sep3] requires that the way these relata evolve is governed by ontic factors outside of the relatum concerned. And **[Sep1]** requires that whatever it is that drives the relata forwards must have some presence in the current state of the universe. This is in tension with the way we used separability above, both in our electron example and in section 4.5.2, and we now explain why.

[Sep1] requires that the original causal nature of this relatum is held in the fabric of the universe and not solely in the causal content of the relatum. That is, the original causal consequent of the relatum in question is made necessary by the configuration and causal content of the whole, or some part, of the universe other than the relatum in question. If we desire to change the way the relatum causes then we must also seek to change some of the relata elsewhere in the universe to effect this new causal consequent. But changing relata elsewhere is forbidden by **[RES4]** and so is not causally permissible. That is, we must conclude that we cannot change the way a single relatum evolves in a deterministic universe without violating other contemporaneous relata.

What is useful to note is that if we were to countenance the alteration of other contemporaneous relata then we would have to invoke more **[Y]** type counterfactuals to effect this. And invoking such counterfactuals pushes the problem back one moment in time. And if we were then to repeat the argument again we would find altering the causal consequent of a separable relatum requires a cascade of alterations moving back in time moment by moment.

In conclusion, we find that there are no time specific counterfactuals that are causally permissible in a deterministic universe. **[Y]** is not causally permissible if the connexions it involves are necessary.

We now turn to hybrid universes. The arguments we have offered stand if we move to consider hybrid universes - ones where some connexions ontically necessitate their consequents and others do not. We ask whether, in such a scenario, a separable relatum - which by its nature must ontically necessitate its future - can be altered: is the altering of separable relata causally permissible?

First, we note that as the connexion here is a necessary one and because the nature of that necessity is not to be found in the originating relatum - by definition of relatum separability - then it must be found in the fabric of other contemporaneous relata or the connexions leading to these relata. As such, to change the separable connexion in this case requires either that other contemporaneous relata be changed [in a controlled manner - controlled so as to effect the desired outcome expressed by **[Y]** type counterfactual] or prior connexions be changed [in a controlled manner]. But either such option violates **[RES4]**: the arguments above apply equally in this case as they did in the deterministic case.

We must conclude that relatum separability, whenever it occurs, is a far weaker notion when considered as an ontic constraint. Altering separable relata or their connexions is never causally permissible.

As a final brief note: it might be argued that **[RES4]** is too restrictive as a condition on causal permissibility and some contemporaneous alteration in relata should be allowed. Whilst this would immediately violate backwards causation - as recall, **[RES4]** was introduced in the context of **[Y]** and the antecedent relatum of **[Y]** is prior to what we seek to change - we might, nevertheless, entertain weakening **[RES4]**. If we do relax **[RES4]** it is still possible to show that earlier relata and connexions will need to be altered when altering the connexion following a separable relatum, and hence violate backward causality, but the arguments would have to be based around the need for *controlled* change rather than on our [newly minted - as **[RES4]** now no longer applies] causal permissibility. Controlled change requires one relatum to be replaced by another specified alternative. The arguments that show controlled change violates backwards causality are both complex and involved so we do not set them out here. We justify this by noting that our discussions above are adequate for the aims of this chapter.

4.5.7 First conclusion: choice is not ontically compatible with determinism

In this section we pause to recap the argument we have advanced above and to state our first conclusion: choices cannot arise from deterministic universes or deterministic aspects of hybrid universes.

We started by holding that we must start our search for choices by considering two counterfactuals **[X]** and **[Y]**. We wanted to understand what it might mean to assert that these were encapsulations of genuine alternative futures when considered at certain times and considered together with the ontic nature of the actual universe.

We then briefly examined Lewis's possible worlds approach to help direct our analysis. Recall that Lewis's approach has a number of features:

[L1] It requires possible, non-actual worlds, to be introduced;

[L2] The past can be considered altered in passing from possible world to possible world;

[L3] We may violate laws locally in passing from world to world.

We rejected Lewis's approach as we are seeking an ontic solution to the problem of choice and this requires us to base our arguments solely within a single causally connected nexus - that is, within one universe alone. Thus we cannot countenance **[L1]**. This prompted us to ask whether it is possible to understand counterfactuals in some manner that is true to the ontic constraints a universe brings with it. We started to investigate this by considering two counterfactuals **[X]** and **[Y]** to see whether the ontic constraints a universe carries with it make such scenarios causally permissible.

We first argued that **[X]** reduces to **[Y]**. We then argued that even **[Y]** cannot be causally permissible if the connexion that joins its parts is necessary - i.e. the relationship between its parts is deterministic. We argued initially, and falsely as an interim step, that **[Y]** appeared to be possible if the relevant relatum is separable. We then moved on to show that this was an argument in error as the idea of ontic necessity and separability are highly restrictive and do not permit the alteration of a single separable connexion without also requiring considerable consequences to past history. And as we ruled out alterations to the past, on the grounds that we reject backwards causation, such changes are not to be causally permissible.

So, at this stage, we must conclude that if we are to seek a solution to the notion of choice through an understanding of counterfactuals whilst staying true to the ontic constraints a universe carries with it, then we must reject choice's involving necessary connexions between relata: necessary connexion of any sort do not permit changes to the structures of their connected relata within a universe; they cannot provide a route to an ontic construal of choice.

In the next section we explain why indeterministic scenario are no better at providing choice. This will then lead us to proposing the idea of immanence.

4.5.8 Second conclusion: choice is not ontically compatible with indeterminism

Here we present two separate arguments. One that shows that considering counterfactuals in indeterministic settings is of no value, and so consideration of **[X]** and **[Y]** are irrelevant. And two, we argue that the ontic choice we seek is not to be found in indeterministic universes. First we argue that any modification is causally permissible in an indeterministic universe but that this then permits any counterfactual of type **[X]** or **[Y]**.

Recall these two counterfactuals required:

> **[X]** P occurs at time t_1 in place of A and causes Q in place of B.
>
> **[Y]** A occurs at time t_1 but causes C to occur instead of B.

If there are no restrictions on how the universe will evolve then both of these are causally permissible. However, so are *any* other counterfactuals and there can be no control at any moment over which alternative transpires as a result of alteration. In Chapter 5, we prove that no conscious being inhabiting a universe can actively direct what happens. As such there is no possibility that human, or other, intervention would counter this statement. In essence we can claim that any version of **[X]** - with P replaced by any relatum and Q replaced by any relatum - can be countenanced, i.e. is causally permissible. Similarly, in **[Y]** any relatum can replace C. However, in both these cases we have too wide a remit for **[X]** and **[Y]**. The essence of counterfactuals is that they are specific. They are specific in that they delineate definitive alternatives to the factual [actual] occurrences and they lock together these alternatives. So P must be locked to Q - as we argued above - and A must be locked to both C and to B in some manner if it is to cause C instead of B - something we have already shown cannot launch in a deterministic setting. Causal permissibility in indeterministic universes says nothing other than 'anything is permitted'. What is important for our arguments, however, is not the loss of constraint that a purely indeterministic universe gives to counterfactuals but rather whether this loss of control is a sufficient base for an ontic construal of choice. The answer is 'no' and it is to this that we now turn.

First the fecundity of alternative versions of **[X]** and **[Y]** that causal permissibility allows is one that derives from their being no presence, held in the contemporaneous fabric of the present, of constraints on the future. If we wish to find alternatives present at some moment in a universe then these alternatives must be enfolded in the very fabric of the universe at that moment. If they are not present and they are not realised then they can be at most conceptual alternatives constructed from a view of the universe and not from the ontic structure in itself. If they are merely conceptual alternatives with no ontic underpinning then we cannot claim they were genuine alternatives. And, as we mentioned briefly at the outset of this chapter, we cannot claim conceptual alternatives represent genuine possibilities on the back of our ability to pilot the universe towards any one out of the set of conceived alternatives. We cannot make this claim as such piloting is not possible and we provide arguments for this in sections 5.6 and 5.7 of Chapter 5.

Our problem, then, is this: alternatives cannot be found in a deterministic universe - as there the fabric of each moment contains the seed for a single and unique future. Nor can they be found in an indeterministic universe as there the fabric of the present contains no alternatives, unless we adopt a form of ontic extravagance for indeterminism - something we touch upon below. Otherwise an indeterministic universe contains nothing at all that indicates either where the future might move to or whether there is to be any future at all. And even hybrid universes cannot provide an ontic construal of alternatives as they suffer in their deterministic and indeterministic parts in the same manner.

What we must conclude here is that choice, as the presentation of genuine alternatives present in the very fabric of the universe, is not to be found in any universe we have examined so far. As a result we propose immanence in the next section.

4.6 Our solution: restricted indeterminism and immanence

Recall, in section 2.10.3, we explained that the ontic laws of nature must be carried along in the fabric of the universe as it evolves. The essentialists, such as Ellis (2007), hold that these laws are carried along essentially in the fundamental kinds that populate the universe. So, for instance, an electron has it in its nature to respond in a certain manner when it finds itself in a certain situation. For the essentialist such as Ellis this is true of all the fundamental kinds that they take to populate the universe. If we reject this form of essentialism, then the laws by which the population of the universe behave must still be present in some manner in the state of the universe at each moment - this was our argument for the necessity of universe essentialism in section 3.9. In the essential universe case the ontic laws must have some actual presence in the universe at each moment so that the state of the universe as it is in itself necessitates the future states uniquely. In this way the future states have some, unrealised, presence in the current states. We might say that some all powerful demon looking in at the universe could, after examining it in full, tell what each and every future state would look like. In this sense, we can say that the future states are **immanent** in the current state of a deterministic universe.

Now, if we turn to an indeterministic universe, one where the future states are not necessitated by the current state, then we have put the case that we no longer have the future immanent in the present in any manner. That is, given the current state of an indeterministic connexion-bound universe, all future states stand as [conceptual] possibilities and there is no state that is immanent in the present. However, the idea of immanence can still be used to give a further [or alternative] form of indeterminism. If we consider that there is no future state whatsoever present in the current state then the evolution of the world can move in any direction or even cease. Alternatively we might nod towards ontic extravagance and say that *all* future states are immanent in the current state but, as the universe evolves, only one comes to fruition whilst the others fall aside as unrealized. And as an intermediate position, we might hold that at each moment a small set of possible futures are immanent in the present and only one of them is realised as the universe evolves and the others then become absorbed, unrealised, into the fabric of the future states.

Here we shall reject ontic extravagance for indeterminism. Instead we take immanence in indeterministic universes to be the ontic presence of a *finite* set of alternative future states in the present fabric of the universe. We reject ontic extravagance here partly for convenience and partly because it is not clear what we might mean by '*all*' in our characterization of indeterminism via immanence by '*all* future states are immanent in the current state.'

An immanent relatum might hold a set of possible futures outcomes within its own content - echoing the idea of essential relata - or the future immanent outcomes of a relatum may be held more broadly in the fabric of the whole universe - echoing the idea of separable relata. This allows us to form two definitions:

DEFINITION 4.2: Essentially immanent relatum

We say that a relatum is **essentially immanent** if it contains within its causal content embryonic versions of a number of possible future outcomes.

DEFINITION 4.3: Separably immanent relatum

We say that a relatum is **separably immanent** if embryonic versions of some possible future outcomes are held in the wider causal content of the universe state it inhabits.

Were we to examine an essential immanent relatum with our ontic magnifying glass we could list all possible future relata that might evolve from it in the causal evolution. An essential [necessary] relatum, in the deterministic sense, is a special case of an essential immanent relatum - but one with only a single immanent future inhering in its fabric.

In fact, if we were to return to our discussions of determinism in Chapters 2 and 3, it is clear that ontic determinism requires future outcomes to be uniquely immanent in current states.

In this study, in Chapters 5 and 6, we shall take the term '**immanent relatum**' to be shorthand for '**essential immanent relatum with multiple futures inhering in it**'. This is because this type of immanence will be sufficient for our needs there.

The idea of immanence for indeterministic systems is not as far-fetched as it might seem. It has some support from physics: if we consider the wave-function of quantum theory and take a realist stance on it, so that we shall hold the wave-function is an actual feature of the world, then we can link it with the idea of immanence.[20] We might consider that the wave-function represents

[20] See discussions in, for instance, Jammer (1974), Bohm & Hiley (1993) and Isham (1995).

$$S_1(t_1) \xrightarrow{\mathbf{c}_a} S_2(t_2) \xrightarrow{\mathbf{c}_b} S_3(t_3)$$

Figure 4.11: Connexions and relata in a simple causal sequence

the actual presence of [some part of] alternative immediate future states of the universe that are sitting unrealised in the present. At some point this immanence becomes [or causes] both a realised outcome and the unrealised ones fall away - usually this occurs as a process of measurement or interaction with other parts of the universe or via a spontaneous change. We do not intend to pursue the details of the possibility of interpreting the wave-function using immanence here as that will take us down lengthy and tortuous side roads. It is sufficient to hint at the similarities in passing.

Immanence provides us with the ontic underpinning we need to claim the truth of our time based counterfactuals without deferring to possible worlds. We shall now unpack this briefly by considering a situation where a relatum evolves into, or causes, a second relatum non-determinately.

Consider the scenario in Figure 4.11. Here we have relatum S_1 at t_1 causing S_2 and then S_2 causing S_3.[21] Now, we ask what in the universe as it is makes it true to say that, at t_1, S_1 could have caused A instead of S_2. If we take it that there are no future relata immanent in S_1 then A does not appear in any form in the actual universe so there appears to be nothing to underpin the statement and make it true beyond our conceiving that A is a possible future of S_1. However, our considerations are not adequate in this case as we seek something about the universe itself and not our deconstruction of it that makes alternative futures real possibilities. If it is the case that A is immanent in S_1 then the possibility of A being the subsequent outcome is a true possibility at t_1 - so long as it is not prevented by other factors - although it drops as a possibility once S_2 is realised. That is, we can now understand what ontic factors allow us to claim **[Y]** type, and hence **[X]** type, counterfactuals are true at appropriate times. The key is to admit that the counterfactuals expressed by **[Y]** are somehow held immanently in the causal content of the relevant relatum even though they do not ever become realised explicitly in the relatum's causal consequent.

4.6.1 Immanence in deterministic universes

As a final note, we look at immanence and determinism. We shall find that this will play a part in the models we set up in the next chapter when we construct our mechanism for free will.

So far we have portrayed an immanent relatum as one that can only be indeterministic. However,

[21] We ignore any causal penumbra.

there are scenarios where an immanent relatum can evolve deterministically. In this section we explain how this arises. Recall that an immanent relatum is one that has inhering in it a set of possibilia. In an indeterministic universe any one of the possibilia inhering in the causal content of the relatum can be realised. However, it is possible to find that a relatum has a number of possibilia inhering in it - in that its causal content viewed in isolation contain only a finite set of possible future outcomes - but also to find the relatum placed in a position where the surrounding universe restricts what it can actually evolve into. In such cases it is quite possible that a relatum can have futures outcomes, A and B say, inhering within it and, at the same time, to have the external universe restrict what can actually be caused by this relatum - for instance, to just the outcome A. In other words, it is possible to have a relatum that is separable in that its future is ontically necessitated by the factors external to it, but also to find that the relatum in question contains a finite set of possibilia inhering within it. As such it is possible to find essentially immanent relata in deterministic universes.

It is the idea of immanence that will provide part of our model for the operation of free will. It is the unpacking of the details of this that we now turn towards in Chapter 5.

Chapter 5

A model for free will

5.1 Introduction

In this chapter we offer a model for free will, drawing on earlier structures and definitions. We move towards our construal of free will in stages. First we set out some key assumptions we shall carry with us and derive some results that will have roles in our arguments. Then we go on to explain how our version of the exercise of free will is to operate.

In outline, then, this chapter runs as follows: we start by setting out three assumptions we shall carry with us throughout the chapter. One assumption is a technical relationship that sets out how the measure structure quantifies a restriction on our observations of the world. The other two assumptions say, in simple terms, that the mental - broadly construed - has a full place in the causal world. That is, we assume that we can muster the techniques of previous chapters when we talk about decisions, choices and the like.
We then set out and discuss three ‘theorems’. Our first theorem is merely a restatement of the assumption that the mental is subsumed under our models. We present it as a theorem - even though we also take it as an assumption - as we offer some discussion to demonstrate that it is *plausible* and not merely a blind assumption made with our final aim in mind. We then present two further theorems which together amount to saying, roughly, that it is impossible for a person to pilot their own thoughts and actions towards a specified outcome.

Then we move to constructing our model for the operation of free will. We start this by setting out a formal definition of a ‘self’, the essence of a responsible agent here. We argue that a self must act ‘consistently’ with its own nature and we explain what we mean by ‘consistently’. We argue that a self must act in this way so that we can attribute responsibility to it for the selections it makes from choices.

We then argue that choices, as alternative futures, are to be found in the nature of what the universe offers to its inhabitants. As such we use the idea of immanence to capture this notion of choices. We then explain why our arguments here lead to a surprising conclusion: we explain why the idea of free will as the ability to do otherwise - that a necessary condition for the exercise of free will is encapsulated in the idea that 'I could have chosen to do A instead of B' - cannot be used as a basis for any ontic conception of free will governed by our models. Finally, we argue that if we were to remove choice, as immanence, from our scenario then the attribution of responsibility would expand beyond the self to encompass such exotica as billiard balls and cheesecake.

We leave the idea of supervenience entirely untouched in this chapter; we do this because we shall demonstrate in Chapter 6 that it does not, in all cases that reduce to causality within a single causal nexus, affect our arguments or central contentions.

5.2 Our assumptions

Throughout this chapter we carry with us three assumptions. In this section we set out these assumptions. In following sections we shall explain the part they will play in our discussions and flesh out some reasons and justifications for holding them.

Informally we can think that our assumptions say two things: first, that we can subsume the mental and the physical entirely under a single causal structure - a structure that can be modelled according to the ideas we have set out in previous chapters; second, that there is a technical requirement which links together the epistemic access we have to the world with the measure structure we introduced in Chapter 3.

We start by stating our three assumptions in broad terms:

> **[A0]** There are technical conditions on the epistemic access we have to the world external to us that are expressible using the measure;
>
> **[A1]** Both 'the mental' and 'the physical' are fully subsumed in a single causal process;[1]
>
> **[A2]** The access we have to the world does not permit us to make *a priori* claims as to causal distinctions between 'the mental' and 'the physical'.

We shall now look at each assumption in turn and flesh out in a little more detail what we take them to say, what part they play in our arguments and outline some general justifications for holding

[1]Implicit as part of this is the assumption that we can apply the measure to both 'the mental' and 'the physical' in the same manner.

to them. We start, in the next section, with **[A0]**. Then, in section 5.4, we rework **[A1]** and **[A2]** into the first of the three theorems we propose. We also offer a discussion as to why we take this reworking - as our Theorem 1 - as a plausible assumption.

5.3 Measure restrictions on epistemic access.

First we deal with **[A0]**. The technical assumptions that it hints at will play a part in some of the proofs and definitions that appear in this chapter. We start by outlining what we require for one relatum to have full epistemic access to another. Such access is the minimal requirement here for any claims that a self - a being subsumed within a causal nexus - has full awareness of other aspects of the universe. That is, in simple terms, we are seeking a condition that will allow us to quantify the epistemic access we have to some part of the universe we inhabit. And in doing so we shall make no assumptions as to the internal mechanisms - such as the way any impingement of the world on a sentient being is conceptualized - or as to the faithfulness of the representation that arises from any such internal processing of the impingement. We shall not worry at all about whether the way we experience and conceptualize the world resembles the world in any sense of the term 'resembles'.

For our discussions, we require only a minimal way of expressing how one part of the world impinges upon another and we shall invoke our measure structure from section 3.5 to help us with this. Recall, we outlined the scope of the measure μ in section 3.5: when applied to a relatum it furnished us with some quantification of the causal content of the relatum, and when applied to a time-ordered pair of relata it furnished us with a measure of the causal influence of one relatum upon the other. As we discussed there, the measure is sensitive to the detailed causal content of a relatum: roughly the more causal content the relatum has the more the value of the measure. In other words, the measure gives us a quantification of the causal content of a relatum and, when applied to ordered pairs, gives us a quantification of how much of the content of one relatum influences, through connexions, the other. As a result, it is natural to assume that, if we wish to obtain complete awareness of a relatum, we must have access to its full structure and hence its influence on us must be quantitatively equivalent to its measured content. In other words, if one relatum, A say, is to have full [epistemic] access to the details of some other relatum, B say, then we require:

$$\mu(B, A) = \mu(B) \tag{5.1}$$

Where (5.1) says that B's contribution to A - given by $\mu(B, A)$ - must be equal to the causal content that B contains - given by $\mu(B)$. Further, we assume that a relatum cannot contribute more than its total content to the future. Reworking equation (3.18), this is expressed as:

$$\mu(\mathcal{S}(t_a), \mathfrak{U}(t_b)) \leq \mu(\mathcal{S}(t_a)) \ \ \forall \mathcal{S}(t_a), \ \ \forall \mathfrak{U}(t_b), \ \ t_b > t_a \tag{5.2}$$

Where this says that the contribution of any relatum $\mathcal{S}(t_a)$ to a future universe state $\mathfrak{U}(t_b)$ - where this contribution is given by $\mu(\mathcal{S}(t_a), \mathfrak{U}(t_b))$ - cannot exceed the causal content of $\mathcal{S}(t_a)$ - given by $\mu(\mathcal{S}(t_a))$. Further, if we were to require global measure preservation, then we could write:

$$\mu(\mathfrak{U}(t_a)) = \mu(\mathfrak{U}(t_b)) \ \ \forall t \tag{5.3}$$

In summary: **[A0]** is taken to assert that we assume our model always obeys (5.2) and we now hold that (5.1) is a condition for complete epistemic access to relata. And further, we have the option to impose (5.3) should we wish.
We note in passing, partial epistemic access occurs when:

$$\mu(B, A) < \mu(B) \tag{5.4}$$

5.4 Prolegomenon to the three theorems

Below we set out our three theorems in detail. Here we give an informal summary of what they say and some indication of how they will be used in the discussions that follow later - discussions that lead us to a construction of a version of free will satisfying the two criteria we set out in Chapter 1.

Theorem 1 says, in simple terms, that we can marshal all the technical apparatus we have set up in previous chapters - connexions, the measure, causality and the like - and apply it to both the physical and the mental within a single model. In other words, causality is all encompassing and we hold that it is the only diachronic ontic relationship that we can use in our analysis. Theorem 1 gives us warrant to talk of thoughts and decisions being causally constrained by the past and contributing to making the future ontically necessary in the same way we can talk about the motion of billiard balls being causally constrained by the past and contributing to making the future ontically necessary.[2] We take it that both the mental and the physical, however construed, are subject to our model and the measure. We promote this assumption to the status of a theorem to emphasize

[2]Of course, if the mental is ever epiphenomenal then its contribution to the future is non-existent.

its importance. Without holding to Theorem 1 we cannot progress as we cannot talk of thoughts leading to actions through a single causal mechanism; we cannot talk of the world causing our perceptions via the same causal mechanism; we cannot talk of determinism and thoughts in the same breath, or under the same single causal model.
What this means for the rest of the chapter is that we can talk of relata and cause and determinism and so forth without regard to whether the relata concerned are 'physical' or 'mental' in the way they might be unpacked in detail.

Theorem 2 argues that no relatum or set of relata can be 'self-piloting'. In other words, no set of relata can 'decide' what its future is to be and then pilot towards this future. In essence we show that any such piloting requires the intervention of separate relata from outside.

Theorem 3 argues that no relatum or set of relata can pilot other relata to some specific outcome. The argument we use is, in essence, an abstraction of one that shows such piloting requires an infinite regress of relata - each separate from the others - and that such a regress cannot be countenanced because it cannot contain a 'first cause'.[3]

5.5 Theorem 1

We start by stating Theorem 1 formally and then we backtrack to our assumptions **[A1]** and **[A2]** to start building a case for the plausibility of the theorem. We shall then briefly discuss one argument that might be used to suggest our claim **[A2]** is invalid and explain why we dismiss it for our models.

Theorem 1 states:

> **Theorem 1:**
> All mental and physical states that interact are subsumed within a single universe considered as closed causal nexus.

Or, in informal terms, this says we can talk of relata and mean such things as billiard balls, mental content and the like. And all relata are subject to the measure and ontically related via the connexion.

[3]A broadly similar version of our argument appears on pages 24 and 25 of Strawson (1984). And another version can be found on pages 30 and 31 of Ryle (1990).

There is no doubt that this claim is difficult to accept at first glance. The mental and the physical tend to sit as very distinct realms within our intuitive everyday view of the world. Indeed Kant takes them as separate realms, even as far as time is concerned. He says,

> But of pure reason we cannot say that the state wherein the will is determined is preceded and itself determined by some other state. For since reason is not itself an appearance, and is not subject to any conditions of sensibility, it follows even as regards its causality there is in it no time-sequence, and that the dynamical law of nature, which determines succession in time in accordance with rules, is not applicable to it. Kant (1964, A553, B581, p.476)

What we aim to achieve in the following discussion is not a definitive rebuttal of the intuitive unease that our theorem might provoke, but rather we endeavour to put a case that the theorem is *plausible*. In the end, whilst we offer discussion, we take it that Theorem 1 is a fundamental assumption of our model as we use it in this chapter.

Recall **[A1]** and **[A2]** from above:

> **[A1]** Both 'the mental' and 'the physical' are fully subsumed in a single causal process;
> **[A2]** The access we have to the world does not permit us to make *a priori* claims as to causal distinctions between 'the mental' and 'the physical'.

[A1] says that everything that we consider within a universe must be related to everything else that is within a universe and that the only mechanism that allows us to relate aspects of a universe together across time is a single version of the causal relationship we outlined in Chapter 3. We can break this down into pieces to make it easier to explain both what is being claimed and why it is being claimed:

> **[A1a]** We have access to a world that is external to us;
> **[A1b]** This access is the origin of our 'perception' of the world;
> **[A1c]** All ontic relationships within the universe are causal and mediated by a connexion;[4] and they are subject to the measure;
> **[A1d]** Our mental content is entirely part of the universe;[5]
> **[A1e]** Our mental content is mediated by causal links;

[4]We ignore the possibility that supervenience is an ontic relationship here. We could be more precise and hold all diachronic ontic relationships are mediated by connexions and restrict any potential supervenient relationships to synchronous relata. We discuss this further in Chapter 6.

[5]We use the term 'universe' loosely here.

We have stated **[A1e]** separately although we can deduce it from **[A1c]** and **[A1d]** if we hold that links across mental content are ontic. We return to this in a moment.

We shall concentrate on discussing **[A1e]** in what follows. First we should attempt to analyse why **[A1e]** might be taken as contentious and then we shall attempt to assuage these worries sufficiently to give some plausibility and weight to **[A1e]** as an assumption.

We shall offer two informal reasons why it *might* be seen as contentious and then explain why we do not take these reasons to be plausible here. Here are our two reasons that we shall argue against:

> **Reason 1**: The mental is clearly a distinct category in the universe from the physical. Our [experience of] mental content is somewhat different from our experience and conception of tables, chairs and the like as solid objects. It seems clear that relationships between tables, chairs, billiard balls and the like might be mediated by connexions and even subject to regular laws but it does not seem plausible that we can extend this connexion-bound mediation, and the corresponding law, to mental content.

> **Reason 2**: We cannot claim that mental content is ontic as it cannot appear as part of an Archimedean perspective. If Laplace's demon were to look in at the machinations of the universe in detail, he would not 'perceive' thoughts and the like. If mental content is not accessible to Laplace's demon, it cannot form part of an Archimedean perspective and hence cannot fall under the idea of being ontic. As such, we can hold on to **[A1c]** and **[A1d]** but not **[A1e]** - as mental relationships are not then taken as ontic.

First we put a case against **Reason 1**. We do not dispute the premise here - that our experiences of the mental are different from our experiences and conceptions of chairs, tables and the like. Certainly experiences divide up in many ways. My experience of blue is different from my experience of toothache. What we do dispute is that we can, from this observation, move to the conclusion that mental content is not mediated by connexions in the same manner that billiard balls colliding are so mediated. We dispute this move partly because we do not admit that differences of experience are sufficient to make such claims. Indeed, it may be that to make such assertions we must rely on either an *a priori* claim that it is legitimate or a mistaken notion of the physical and causality. We dismiss any such *a priori* claims below when we discuss **[A2]**. Here we concentrate on discussing how a mistaken notion of causality and the physical might lead to someone's advancing **Reason 1**.

First, if we deny that we have some *a priori* legitimacy to claiming the physical and the mental are fundamentally distinct realms - in that one is subject to connexion-bound causal mechanisms whilst the other is not - then if we are to uphold the gist of **Reason 1**, we must seek some justification to move from the variation in experience and conception to the assertion that we might suspect that the mental cannot be subject to the connexion-bound causal mechanism. Below, in section 5.5.1, we shall put the case that it is consistent for the two realms to appear to be governed by different sets of rules - or no rules at all - but still to be both subsumed under a single causal nexus. This suggests that the variation in experience is consistent with the assertion that both realms are subsumed under a single causal nexus.

So if we are to seek to defend **Reason 1**, we cannot do so on the basis that a causal nexus cannot accommodate differing types of sub-nexus. That leaves us only two options when trying to justify **Reason 1**: we must either seek a justification of it in the nature of the connexion itself - by explaining how a connexion-bound mechanism cannot possibly apply to mental content; or we must seek to explain the distinction by looking back to the manner in which we access mental and physical aspects and how such access might lead to claims they are distinct realms. We deal with the second case both when discussing **Reason 2** and in section 5.5.1 below. So here we shall concentrate on revisiting the causal mechanism we set out in Chapter 3 and explaining why we *cannot* dismiss the possibility that they apply to mental content. This then blocks justifications for **Reason 1** based on the nature of the connexion.

In Chapter 3 we set out a mechanism in general terms that we took to be fundamental to all relationships between non-synchronous aspects of a universe. In fact, we went as far as to define a universe in causal terms. And further, we deliberately left the details of the connexion - what we meant by transfer or influence from moment to moment - slightly vague. It is this latitude with the nature of causal content and connexions that will now be invoked. If we take the idea that the connexion is a reflection of mediation by something such as energy or energy-momentum then it is certainly difficult to grasp how something such as mental content can be mediated by it. For instance, it is hard to conceive of how we might deconstruct the movements of reasoning in terms of energy-momentum. However, we should notice that we deliberately did *not* settle on one notion of mediation over another and we did not settle on one that has physical overtones - such as energy or energy-momentum or some such related notion. We did not settle on such notions because they carry with them implicit ideas that causality is a physical phenomenon and hence inject into the idea of causality more than we wish to place there - by making an assumption that the physical is a legitimate category and ontically distinct from the mental. However, it is quite possible that the connexion-bound mechanism involves something altogether different from our notions of energy,

energy-momentum and the like. We might, for instance, propose that the connexion and the measure concern 'information'. It does not seem so far fetched to consider that the mental is a process of information flow in some manner.[6]

The point is that it would seem that a narrow physical construal of causality, and causal content, in terms of a physically-couched idea could lead to the dismissal of claims that mental content is itself subject to the causal mechanism. And to construe causality in such physical terms is to load the final conclusion into the initial assumptions. If, however, we consider that we might think in broader terms then the idea that the mental is subject to the causal mechanisms seems less far-fetched.

We now turn to discuss **Reason 2**. This claims, in essence, that if we consider the ontic to be delineated by what can be observed within the universe by some Laplacian demon placed outside of the universe - and hence one with an Archimedean perspective - then it seems quite legitimate to argue that, as this demon cannot possibly 'perceive' the mental content that we claim inheres within the universe, then we cannot legitimately make the claim that this ontic view includes the mental. Our arguments against this hark back to the discussion we set out following our initial examination of Laplace's quote on determinism in section 2.2. There we made the distinction between considering Laplace's demon as a convenient way of talking/thinking and the need to be careful when talking in terms of a demon's perspective: we must be careful not to take notions carelessly across the boundary between internal perspectives - those that we have as beings subsumed in the universe - and Archimedean perspectives - ones that are attempts to model and explain the universe from a perspective outside of it. We claim that it is precisely an error of this sort that leads to the claim that Laplace's demon cannot 'perceive' mental content and so mental content cannot be subsumed under our causal mechanisms. We have no issue with holding that we, as beings subsumed within the universe, have distinct internal and external viewpoints - so, for instance, I can know [in theory] all about another's brain and environment from an external point of view but still fail to 'perceive' their mental content. But we also hold that we cannot take this analogy across to the demon's viewpoint unless we already have adopted the assumption, as part of the analogy, that the mental is somehow outside the causal nexus in the same manner in which the demon is outside of the universe. Rather, we take it that the demon's viewpoint is such that it can discern all the details of anything that is part of the universe - considered as a closed causal nexus - and that this *includes* the mental content of beings within the universe. That is, we hold the analogy of external and internal viewpoints is valid for beings within a universe but need not

[6]The *loci classici* of information theory are Shannon (1948a) and Shannon (1948b), but here we do do not delineate too closely what 'information', in *our* terms, might be.

translate across to the viewpoint we claim for the demon. The mode of 'perceiving' of an external demon is not analogous to the mode of perceiving of a human bound within a universe.

5.5.1 Looking at [A2]

We now turn to consider **[A2]**. We take it here that we cannot ever justify *a priori* claims but, if we are to hold them, then we must hold them as axioms of a model. And here we simply choose not to make the *a priori* ontic causal claim.

We might ask why we wish to make the claim **[A2]**. We do so because we wish to be able to argue for **[A1]** and if we deny **[A2]** then it gives us a route to deny **[A1]**. If we deny **[A2]** we could argue that our *a priori* access permits us to claim that the mental and the physical are distinct realms and do not interact *causally* within a single connexion-bound mechanism. Recall why we want to hold to **[A1]**: we want to hold **[A1]** because we want to be able to apply arguments based around determinism and indeterminism and the models we have set up to the issue of free will. If we cannot apply determinism and indeterminism in any form to the mental, or its relationship to other aspects of the world, then it is hard to unpack how we might take, or argue for, any position on free will. It is primarily for this reason that we hold to both **[A1]** and **[A2]** in this study.

There is one final comment we shall make. This concerns an argument, or a way of thinking, that might be invoked to counter **[A2]**. We put a loose version of this argument below and then explain why, within the bounds of what we have set up over previous chapters, the argument is not strong enough to counter **[A2]** or Theorem 1.
We start by setting out the argument we wish to discredit. This argument proceeds as follows:

> **[Arg 1]** I have some general intuitive notion of what constitutes the mental and what constitutes the physical. This intuition is based on my experiences;
> **[Arg 2]** I notice that the relationships between the physical aspects are of a distinct sort from those amongst the mental aspects;
> **[Arg 3]** This suggest that I can classify the mental and the physical by different sets of law-like relationships: the physical is governed by one set of relationships - which I shall call p-relationships. The mental is governed by other distinct m-relationships;
> **[Arg 4]** I can now use these two distinct sets of relationships to distinguish between the 'causal realm' of the mental and the 'causal realm' of the physical;
> **[Arg 5]** The 'causal realms' of the mental and the physical must be separate realms: they are not subsumed under a single causal nexus.

Certainly it is true that the mental realm and the physical realm *seem* to be very different. My toothache is certainly somewhat different from my experience of a table and the solidity, spatiality and so forth of the object I associate with this experience. As such we can accept **[Arg 1]** to **[Arg 3]**. That is, we can accept that the relationships within each realm are of a different sort - a different sort in the way they unpack as theory. For instance, we might find that the physical realm is governed by mathematically-couched laws but the mental realm is not. However, a move from **[Arg 3]** to **[Arg 4]** is not so simple. If we accept the argument of Chapter 2 that there need be no easy move from epistemic theory to assertions about the web of connexions, then **[Arg 4]** is not so straightforward. For sake of argument we shall allow the assertion of **[Arg 4]** to follow from **[Arg 3]** as we intend to concentrate on rebutting the move from **[Arg 4]** to **[Arg 5]**. Here, then, the move from **[Arg 4]** to **[Arg 5]** is what we are to challenge. We challenge it on the ground that **[Arg 5]** does not follow from **[Arg 4]**. And of course this does not mean that we shall deny that it *might* be the case that the mental and the physical - construed via **[Arg 1]** to **[Arg 3]** - are indeed entirely distinct realms and not both subsumed under a single causal structure. But, as we shall argue, it does not follow from the earlier statements so must be asserted, if it is to be asserted at all, *a priori*. Why, then, do we reject the move to **[Arg 5]**? We reject the move to **[Arg 5]** because there is nothing in what we have presented so far that forbids a causal nexus from dividing into sub-nexuses governed by different patterns of behaviours or even possessing no such finite pattern at all. We can accommodate distinct causal realms - distinct in terms of the patterns they exhibit - within our overall causal nexus. And if we can accommodate such realms within one scheme then we cannot assert that **[Arg 5]** follows from the **[Arg 1]** to **[Arg 4]**.

We note that if we were to accept that **[Arg 5]** is correct - for instance, if we take it as *a priori* - it leaves us with a difficulty which we hinted at in Chapter 2 when we considered Humean universes. There we mentioned that the lack of a connexion, or ontic relationship, between the ‘objects’ of a Humean universe makes it hard to claim that the universe impinges on us in some manner. That is, if we do maintain that the mental realm and the physical realm are of entirely different natures then we must ask: What is the nature of the interface between them? How do we gain knowledge of the world and use it in our mental machinations? That is, asserting **[Arg 5]** is not an innocuous move and carries difficulties of its own.

5.6 Theorem 2

Theorem 2:

No relatum can self-pilot towards a given outcome.

In this theorem we put the case, drawing on work of previous chapters, that no relatum can self-pilot. We say that a relatum can self-pilot if it can direct itself to evolve into or cause any one out of a number of possible futures.

We show self-piloting is not possible by examining the three causal types of relata we have met in previous chapters:

Relata that cause their future necessarily.

Relata that cause their future via pure indeterministic connexions.

Relata that are immanent relata.

From the discussion in Chapter 4, it is clear that only immanent relata have the possibility of self-piloting towards alternative futures. Necessary relata must evolve [or cause, or become] according to their nature - and so do not have a 'choice' of futures to evolve into. Purely indeterministic relata evolve serendipitously and there is nothing within them, according to their definition, that can lead them to any one specific outcome.[7] And in contrast, an immanent relatum has a number of possible distinct future outcomes it can evolve into. The question is whether it can pilot itself into one of these rather than another. There is nothing we have yet placed in the nature of an immanent relatum that will permit it to direct itself; rather, in what we have proposed, what it evolves into must either set by serendipity or set by outside influences - such as relata external to it - impinging on it in some manner and directing it to one outcome over another.

We must then ask if we can modify our understanding of immanent relata in some way to permit them to self-pilot. We might suggest that one part of an immanent relatum can move to modify another part of it in such a way that it is piloted towards a desired goal. However, to do this that part which is initiating the changes elsewhere within the relatum must itself be able to pilot towards effecting the desired modification of the other parts. And to do this it must require that some part of this relatum be able to pilot the rest of it into the appropriate outcome. This clearly leads to a regress which cannot end. As such no individual relatum can pilot itself.

In informal terms our arguments says two things. First, to pilot to some outcome a relatum must be able to take the outcome that is desired and then modify itself so as to ensure that is the outcome that will ensue. And then we see that to modify itself towards an outcome means it must somehow seek to adjust the future by modifying the present. But it must modify from a standing position which is fixed. And if something is both fixed and required to change at the same moment we find we are at an impasse. The problem of the infinite regress then encapsulates the impossibility of

[7]We take the view here that indeterminism is not akin to an ontically extravagant form of immanence.

modifying the fixed present.

We can also look at this theorem as a consequence of our discussions in Chapter 4. There we argued that we cannot modify the present in any manner to give alternative futures without modifying the past before it. Here we are faced with a relatum at time t_0 and require it to be able to select which future it is to evolve into. We know from our arguments in Chapter 4 it cannot alter its past to alter its present to guide its future. And it cannot change its present nature at t_0 as that is fixed. And so it cannot at t_0 act to move to a different specified future in a controlled manner. Further, our idea of immanence allowed for alternative futures to inhere in the present. If this immanence can be modified internally at t_0 to give only one inhering future, then the relatum in question can no longer be an immanent one at t_0 and so must have been able to change from one specific relatum to another entirely different one without moving forward in time. Our causal mechanism forbids non-temporal changes such as this.[8]

5.7 Theorem 3

> **Theorem 3:**
>
> No relatum can pilot another relatum to a given outcome.

This theorem demonstrates that no relatum can pilot another relatum to any specified outcome from a set of possible outcomes. That is, no relatum can direct another relatum to evolve into, or cause, any particular specified outcome from a set of possible futures.

It is clear from the discussion in Chapter 4 that the directed relatum must be an immanent one. Now, consider an immanent relatum **I** with two possible outcomes inhering in it - we call these outcomes **P** and **Q**. Consider some other relatum **D** which is to direct **I** into one of these two outcomes - perhaps, say, **D** must pilot **I** to become **P**. That is, **D** must itself ensure it is in a state to cause the transition **I** to **P**. If it is already in the appropriate outcome then it will cause the transition but do so without any need for direction. So, for sake of argument, we must assume that undirected **D** would pilot **I** to cause **Q** and so it is then incumbent on **D** to self-pilot in such manner that it would in fact cause the transition **I** to **P** and not **I** to **Q**. However, as we argued in Theorem 2, **D** cannot self-pilot in such a manner. So rather we must either admit defeat or assume that **D** is itself an immanent relatum, distinct from **I**, that can be directed into the correct outcome - to pilot the transition **I** to **P** by some further relatum **F**. The argument must then repeat itself so we find that **F**, in turn, requires some **G** to pilot it into the correct result to ensure that

[8] We do not pursue arguments that involve backwards causation and causal loops here.

D is piloted; and then **G** requires an **H** and so on. The regress is infinite and there can be no first relatum that initiates the chain. As such we can conclude no relatum can pilot another relatum into some specified outcome if that specified outcome is against the nature of the relatum involved.

It is useful to note that the arguments here apply even if the desired futures are incompletely specified. For instance the desire might be to pick up a tea cup resting on a table. The exact details of how this cup is to be lifted can remain vague. Even then the possibility of self-piloting to such an outcome cannot be countenanced as some infinite regress will be needed and this, as we argued above, blocks self-piloting as it must lack an initiating cause.

We now turn to setting out and explaining how we shall achieve a version of the operation of free will within our model.

5.8 The consistent self

A consistent self is to be the minimal ontic model of a being in the universe who is capable of acting in a manner that is responsible. We take a self to be a persisting form and one that can be said to be maximally responsible for its own nature at each moment. We achieve the persistence of form through a definition that links the self at each moment though causality; and we achieve the self's responsibility for its own nature though evolutionary consistency. We leave a discussion of why our definition leads to responsibility until later.

In this section, then, we set out to explain what we mean by a 'consistent self' without worrying at all about what additional necessary conditions there may be on the constitution of this self. For instance, we ignore here whether the self we define must additionally possess a consciousness or a self-awareness - whatever the terms may mean. We seek only to set out the necessary outer boundaries of a self in order to explain, in the subsequent sections, how it may be considered to evolve 'consistently' with its own nature at any time and how it is to interact with immanent possibilities. Any object in our model - whether it be a conscious being, a billiard ball, a star - must be constituted of concatenations of relata that are themselves causally involved. By the definition of universe we offered in section 3.7, everything that interacts causally constitutes aspects of the universe. Additionally, Theorem 1 above suggests that we need not worry, when talking of objects and the like, what types of relata constitute them. So, the first step in constructing a persisting object which we shall call a 'self' is to delineate it as a collection of time-indexed sets of fundamental relata that are causally related. Our task is then to set the boundaries of the causal relationship between aspects of this self at different times and justify the boundaries we construct in terms of the results we aspire to achieve.

We start by defining a momentary-self, $\Xi_x(t_a)$, at a given moment t_a as a well-defined [structured] set of fundamental relata at that time. We denote this either as a set of fundamental relata or a single relatum that constitutes the set:

$$\Xi_x(t_a) = \{f\mathcal{S}_i(t_a); i \in \mathcal{I}\} = \mathcal{S}_x(t_a) \tag{5.5}$$

Where the subscript on Ξ labels a specific self and $\mathcal{I}$ is an indexing set.

And so the relata that make up a self *in toto* constitute of a time-bound set of these momentary selves, the time-bound delineating the start and end of the object's existence:

$$\Xi_x(t_\alpha, t_\omega) = \{\{f\mathcal{S}_i(t)\}, \forall t : t_\alpha \leq t \leq t_\omega\} = \{\mathcal{S}_x(t), \forall t : t_\alpha \leq t \leq t_\omega\} \tag{5.6}$$

To complete our initial specification of a self we must explain how the momentary selves at different times are to be related together. We shall give only necessary conditions and explain briefly why sufficiency conditions are not easy to outline.

We shall require that each fundamental relatum that constitutes part of momentary self at any given time has, as part or all of its immediate antecedent cause, fundamental relata that are part of the previous momentary self. And further we require that the causal connexion between these relata is a necessary one, for reasons we shall explain below. Formally, the first part of our condition states:

$$\forall f\mathcal{S}(t_1) \in \Xi_x(t_1) : \mathfrak{C}(t_1, t_0)(f\mathcal{S}(t_1)) \cap \Xi_x(t_0) \neq \emptyset \tag{5.7}$$

This is the requirement that part or all of the current self participates causally in every aspect of the next momentary self.

As mentioned, it is also a requirement that this is a necessary participation mediated by a necessary connexion - so that all the causal relationships between elements of a self at different times are ontically necessary ones. What we do *not* require is that the future momentary self is *solely* caused by the current momentary self. This is because the self can be changed and modified by outside influences too. For instance, I might be hit by a boulder whilst walking along the road, or I might have a new idea stemming from something I have read in a book, or I might be affected by a beautiful painting. The condition of ontic necessity is necessary but not sufficient as outside influences have their part to play in shaping the momentary self. The influence from outside might then make the exact delineation of a self at each moment difficult, in a deterministic universe, and probably impossible in a non-deterministic one. Our aim is only to delineate a self adequately to facilitate examining the idea of a self making choices that are both free and consistent with its own nature at the moment of choosing. And that it is maximally responsible for its own nature from moment to moment. It is to the idea of consistency we next turn.

First, the usual notion of consistency is one that derives from the idea of the absence of contradiction. And in turn this tends to have roots in a formal framework of some kind. In Chapter 2 we explained that predictability, determinism and indeterminism come apart in our ontic models. As such if we were to appeal to some rule-bound structure for our notion of consistency we run both the risk that the rules do not reflect ontic factors and we also run the risk that there may be no rules to rely upon anyway. As such, we cannot attempt to construct a definition of consistency using the epistemic laws of nature if we are to make consistency ontically robust.

Rather here we need to define consistency in ontic terms. And recall that eventually we are going to use our notion of consistency as the basis for attributing responsibility. As such we want our consistency to associate future relata clearly and unambiguously with some current relata and to do so in a way that makes the association in some sense necessary. We shall explain later, in section 5.10, why we want the association to be necessary. The first condition - that we can associate future relata clearly and unambiguously with some current relata - is one that the causal relation gives us. The other condition - of necessity - can only be ontic necessity here as we have already dismissed the possibility of using necessity gleaned from some formal articulable theory. So, we shall base our notion of consistency on ontic necessity.

In essence, we shall say that some set of relata B is [ontically] consistent with a preceding set of relata A if and only if every fundamental relatum in B has some fundamental relatum or relata in A as part of its cause and, further, that the causal linkages between these fundamental relata are all necessary. However, this is not to say that A ontically necessitates B as we leave open the possibility that other relata beyond A also have causal influence on B. We can see that our specifications above lead immediately to the momentary self being [ontically] consistent with its preceding momentary self.[9]

Informally, we can think that the motivation behind this definition of consistency is to ensure that we have the self acting and evolving in a manner that is, at least in part, fully compatible with its nature at each moment - that is achieved via the causal aspect of the notion of consistency; and in a manner that makes the self at one moment partly responsible for the self at the next mo-

[9]It is important to note that ‘consistent’ here is a technical term. Issues of acting consistently can only arise in deterministic subsystems; it is not meaningful to talk of systems acting ‘inconsistently’ with their nature.

ment - and that is what is achieved by requiring the causal links to be ontically necessary ones. We explain, in section 5.10, why we take it that ontic necessity gives us purchase on responsibility.

There is an additional comment that is useful to make at this juncture. The self as we have defined it can exist in both deterministic and indeterministic worlds, and there are some important consequences to this that we shall meet below. What is important to emphasize here is that any such self, even if it exists in a universe that is amenable to prediction, does not itself have to be amenable to prediction when considered in isolation. This means that our definition of a self does not have to be considered as nothing more than a glorified machine running along predefined algorithmic tracks. In fact there can be no logical jump from our definition of a self to assertions that we are, in any sense, an algorithm-governed machine. This does mean that we are dismissing the possibility that a self could be mimicked by some form of advanced computational device. That stands as a possibility independently of our assertions here.

We now turn to explain how the notion of a consistent self, as a crude model for a human being, can be combined with the idea of immanence to produce some notion of free will.

5.9 The machinery for the exercise of free will

In this section we set out what we shall take to be the mechanism for the exercise of free will but we shall postpone our arguments as to why this is an optimal mechanism and how it achieves the two aims we set for free will in Chapter 1 until the following sections.
Recall our two conditions from Chapter 1:

> **1.** The universe presents genuine choices, genuine alternative futures, to its inhabitants;
> **2.** The inhabitants 'choose' between these choices in such a manner that they are clearly and solely responsible for their choices, their selections.

What these say is that when a choice is made by a being we shall require that the outcome of the choice can be attributed to the being in some manner that gives us warrant to bestow [sole] responsibility for the choice on the being; and additionally we require that the choice that is presented to the being is a genuine one and not an ersatz one.
We shall require two ingredients for our mechanism: one is the existence of a self that evolves, and hence acts and makes choice, in a manner that is consistent with its nature at the time of acting or choosing. Recall that above we stated that a consistent evolution of a self is rooted in the determ-

inistic, and hence [ontically] necessary, nature of its evolution. The second ingredient we require is that the choices presented to the self are genuine and not ersatz and that they are encountered serendipitously. As we discussed in the previous chapter, genuine choices constitute an immanent set of possibilia inhering in some relatum.

Our mechanism for the operation of free will constitutes two structural requirements:

> **[FW1]** A self that evolves consistently;
>
> **[FW2]** Choices are presented as immanent possibilia inhering in relata and encountered serendipitously.

There is a third element we require for the exercise of free will and this involves a specific restriction on the combination of a consistent self and an immanent relatum. We shall require that the outcome of any encounter between an immanent relatum and a consistent self is always fully determined, and hence ontically necessitated by, the combination:

> **[FW3]** The outcome from an interaction between a consistent self and an immanent relatum is ontically necessitated by the consistent self at the time of the encounter.

In section 5.10 below, we explain why **[FW3]** is the means through which we justify the attribution of responsibility to the self for the interaction.

We leave no room for alternative possibilities to such an encounter. Once an immanent relatum encounters a consistent self, the causal future of the immanent relatum is sealed.

In this model of the exercise of free will that we are proposing it must be emphasized that the nature of the self at each moment cannot be said to be the same as at previous moments. There may be much in common from one moment to the next, as befits our experiences that we are fairly much the same from moment to moment, but the self is changed and modified by impingements from outside of itself.[10] So, for instance, I may be hit in the head by a rock and this may alter me in some way.

Let us examine how this mechanism pans out by means of an example. Imagine a simple situation where I am presented with the possibilities of turning left or turning right at a junction. I choose to turn left. Now the choice to turn left or right at the junction is presented to me by the external world [as an immanent relatum, we assume] and the choice is made by an interaction of the presentation of the alternatives and the internal machinations of my thoughts, and then movement of my limbs etc follow.

If I were to reach the junction at a slightly different time then there is no guarantee the same

[10] This is not an argument but a comment. It may, of course, be the case that we are radically different from moment to moment when viewed from an Archimedean perspective, but remain seemingly similar when viewed from an internal perspective. Certainly we assume that the two perspectives are somewhat synchronized but we can only ever take this an assumption.

decision would be made. The combination of a consistent self and immanent relatum that we take to lead to a specific outcome, determined by the nature of the self, must be time specific and depend upon the exact nature of the self at the moment of encounter between the immanent relatum and the self.

There is, of course, also the possibility that the self encounters outside impingements which are not immanent relata. For instance, I may find myself struck by a rock rolling down a hillside and this will then alter the manner in which I, as a self, evolve. It may be possible that the self encounters truly indeterministic relata, that are not constituted of immanent parts, and then the outcome of the evolution of the self in combination with these would itself be indeterminate. In the first case, where a rock falls upon my head, we can assume that the rock's course is fully determined and so the future relata of the self are themselves still fully determined by the self in combination with these external deterministic impingements. If the self does encounter truly indeterministic relata and finds itself no longer deterministically determined at some junctures in its evolution then the future self is determined only in part by the current self, not fully. This latter case should not matter unduly to grasping a view of the integrity of a self as even the way the self encounters immanent choices is not determined in advance. This is because the encounter between the self and an immanent relatum must be by chance in accordance with **[FW2]**. We only require that the self determines the results of its encounters with immanent relata and we suggest that the bulk of what we would construe as active choices are in fact ones that fit this pattern.

The question remains as to why we can take this model as being one that offers an optimal analysis of the exercise of free will - one that satisfies our two criteria. It is to this we now turn.

5.10 Towards a notion of free will: responsibility

In this section we explain how the machinery we outlined in section 5.9 above allows us to attribute responsibility to a self for its choices. We postpone a discussion of why they allow us to attribute *sole* responsibility until the end of section 5.11.

We start by explaining why Theorems 2 and 3 can be used to demonstrate that a counterfactual ontic construal of responsibility is invalid. This result will then lead us to an explanation of why we have linked our notion of consistency with ontic necessity.

Our dismissal of counterfactual construals of free will might be seen both as surprising and counter-intuitive. It is surprising as counterfactual scenarios tend to feature in versions of free will. And it is counter-intuitive as we certainly *feel* as if we are piloting our deliberations from the moment the choices are presented to making a selection and that we could have deliberated other than we did: I chose apple pie but I could have chosen cheesecake. Here we might interpret this truth as stating that in the moments leading to choosing we pilot ourselves to selecting one rather than the other of some presented choices but, given the same set up again, we could easily pilot ourselves towards a different choice.

We shall put the case that such counterfactual construal can have no place in an ontic unpacking of free will. Often the idea of free will is encapsulated in the truth of counterfactual scenarios such as the following:

[FACT] At time t, I chose A

And I exercise free will at time t if there was also at least some other choice open to me, B say, where B is independent of A, such that the following is true:

[SUPP] At time t, I could have chosen B.

So the truth of **[SUPP]** is often taken as a mark of free will. I am free to choose if I could have piloted my thoughts to choices [and then actions] other than those choices that I actually selected. And in piloting my own thoughts [and actions] in this manner I am responsible for them and so can be deemed responsible for my decisions and actions.
We shall put aside the issues with counterfactual and choice that we took up in Chapter 4, and instead take a more intuitive look at **[SUPP]** to explain why it cannot be a legitimate encapsulation of the route to the attribution of responsibility for the choice made.
Before we explain why this is the case, we note that often it is taken that this encapsulation may fail because in a deterministic universe there is no possibility of an alternative choice B, so **[SUPP]** fails. And, in an indeterministic one, the choice itself would be undetermined because of the lack of control over events and so the idea of responsibility for the choice fails as it is rooted in serendipity not purposeful deliberation or some such. Our mechanism succumbs to neither of these objections.
In contrasts to these objections to **[SUPP]** our notion of free will outlined in section 5.9 above requires that the exercise of responsibility must be antithetical to the possibility of statement **[SUPP]** being true. That is, our mechanism requires **[SUPP]** to be false if responsibility is to be attributable. Why is this? If it were the case that at some time I am presented with immanent choices then Theorem 3 forbids the final selection deriving from me through self-piloting. If there is no self-piloting then, in the absence of any other mechanisms, the selection between the presented choices must come down to serendipity. In this case then, the path the combination of the self and the [immanent] presented choices take is not determined by the nature of the self. If it is not

determined by the nature of the self then we cannot then attribute responsibility to the self for the resulting choice - there is nothing about the involvement of the self, in the absence of self-piloting, that necessitates one outcome over another. That is, in the absence of self-piloting **[SUPP]** cannot encapsulate responsibility. If **[SUPP]** is true it must be as a result of the indeterminate nature of an interaction between a self and immanent relata and this cannot then lead to attributions of responsibility.

So if a mechanism is to allow the possibility of responsibility then we must dismiss characterizations of free will in terms of **[SUPP]** *ab initio*. Hence, we must reject deliberation-dependent counterfactual encapsulations of the exercise of free will in all cases: they cannot be the correct encapsulation if some form of responsibility is to be part of our ontic understanding of free will. And we see the reason for this is the intervention of Theorem 3.

How, then, does our mechanism allow for the attribution of responsibility for the choices made by a self?

Our mechanism is clear: genuine choices, as immanent relata, are present in a universe and if they encounter a consistent self then the result is fully determined by the nature of the self at the moment of the encounter.

We claim that it is because the outcome is determined by the nature of the self at that moment that we can attribute responsibility to the self for the outcome. That is we claim responsibility can only be taken to be derived from a mechanism such as the one we have set up in section 5.9 above. That is we hold that:

> **Responsibility for selections from choices by a consistent self *must* be derived directly from the deterministic selection between immanent relata inhering in the present *and* the serendipitous encounter of an immanent relatum with the self.**

In other words, we link together determinism, i.e. ontic necessity, with responsibility. And we also require that the self encounter the choices from which it selects serendipitously. Here we largely set aside the second aspect of our requirement - that of serendipitous encounter - and we return to it in section 5.11 below. We now turn to explain why we have defined responsibility using ontic necessity. That is, we explain why our notion of consistency is part of the key to responsibility here.

In essence to attribute responsibility we want to demonstrate that the responsible agent gave rise to some choice and that they did so in a manner that allows us to state the specific choice was made solely because of the participation of the responsible agent - later we shall explain that the 'solely' links with the serendipitous encounter. We have already explained above that the causal aspects of our mechanism allow us to state that the agent gave rise to the choice so we concentrate on how our mechanism shows us that the choice was made the way it was because of the participation of the responsible agent. As we are looking for an ontic construal - and recall we can have no recourse to theory - we have only two ways to turn. We can either use indeterministic causation to justify the claim that the choice was made as it was because of the participation of the responsible agent; or we can use deterministic causality. It is clear that indeterministic causality does not link the result of a selection to the presence of a self in our mechanism. If the self participates indeterministically in the selection then the immanent relatum can give any one of a number of results and the self's presence is not an essential part of this. If, however, we take it that the self causes deterministically then the outcome is clearly the direct result of the presence of the self. It is this fact that allows us to state the self is responsible for the choices it makes via our mechanism.

Informally, we can see that this says, in effect, a self is responsible for its selections when it selects in accordance with its nature. And the only way to have a specific nature to act in accordance with is for the self to cause outcomes deterministically. We can then extend this to the self's evolution and say that the self is partly, and maximally, responsible for the way it is at each moment.

Our mechanism, as it stands, still has an uncomfortable air: it fails clearly to distinguish between a billiard ball and a person in the exercise of free will. Our problem is that our definition of a self, eschewing all explicit reference to mental aspects, fits a billiard ball moving along deterministically. If it is the case that a billiard ball encountering an immanent relatum then determines the outcome for the encounter, as we said a self must if it is to exercise free will, then we have grounds to bestow responsibility onto billiard balls as well as onto people, *qua* consistent selves. This is an uncomfortable situation and begs for some suggestion as to how it might be resolved or some arguments as to why it cannot be resolved.

First, the most obvious ploy is to claim that as Theorem 1 gives us warrant to talk of mental aspects themselves as relata or aspects of relata then we could perhaps only allow responsibility to be bestowed upon such encounters of a self and an immanent relatum when part of the process that leads from the encounter to the action is provided by relata that have some specific type of content, namely mental. The invocation of this sort of mental content for specific relata is not ruled out by Theorem 1 and this would seem a legitimate move. So we might then decide to modify our definition of a consistent self to include mental relata, or at least relata with mental content, specifically mentioned. An alternative would be to work the other way around. That is to say that it is only a certain sort of deterministic conglomeration of relata - satisfying our definition of a consistent self - can interact with immanent relata to produce determinate outcomes. And we could then postulate that such conglomerations are precisely those that possess, in some sense, mental content.

There are two further possibilities: we could countenance, when considering billiard balls and people, that we can attribute responsibility to both without discrimination. The other is to require the presence of some mental aspect for our mechanism to work but then allow billiard balls to possess mental aspects. Neither of these is particularly appealing. If we were to decide upon one, we would suggest that mental aspects play a crucial role in the process. That is, we would hold that the presence of mental content, in some form, is necessary for the working of our mechanism. We leave this suggestion and the problem we have outlined unresolved here.

Our model steers a path between determinism and indeterminism, drawing on determinism for the operation of the self to allow the ascription of responsibility, and drawing on a type of indeterminism to keep real choice in the mix. However, we have yet to explain why immanence plays such a crucial role. We ask: if we were to remove immanence, then might we still be able to attribute responsibility for choice and action to a self evolving consistently? Indeed, it appears that the responsibility we have bestowed upon our consistently evolving self derives from its deterministic leaning and not from the choices it encounters. It is now to this issue we turn: to explain why the possibility of immanence both deserves the soubriquet ‘genuine choices’, i.e. alternative futures, and constitutes an essential ingredient of our mechanism.

5.11 Towards a notion of free will: immanence and choice

In this section we first make some comments to justify our claim that immanence is what provides choices in our model. We then go on to explain why immanence is central to our model. In the absence of immanence there can be no claim that we have free will.

First, some comments on immanence as providing the notion of choices. We have already argued in the previous chapter that immanence is a method of introducing an ontic underpinning to the notion of choices construed as the possibility of alternative futures - a possibility that inheres ontically in the present. We justify the name ‘choices’ for immanent relata when encountering consistent selves by claiming that a notion of choice must have two facets:

> **[Choice1]** Choices must indicate future possibilia inhering in the current universe;
>
> **[Choice2]** The choices must be presented to a self and an outcome must be selected so as to permit the attribution of responsibility for the choice taken.

[Choice1] might seem to allow conceptual possibilities as choices. It might seem that I can conceive of two possible future actions for myself - eating cheesecake or apple pie - and that prior to my deliberations both are genuine possible paths for me. They are both genuine possibilities because I can deliberate to either path and I have the capacity to act upon my deliberations. This amounts, in part, to the claim that conceptual alternatives are immanent in relata that are aspects of my mental states and hence genuine choices; and that my deliberation leads from these immanent conceptual states to selection. However this cannot be countenanced in our model: we have explained in Theorem 3 that deliberation of the sort required here is not a possibility if we are to maintain responsibility.

As such, we must take it that **[Choice1]** must allow for the ontic presence of future possible paths independently of a self. It is then **[Choice2]** that allows these immanent relata to be called ‘choices’ when they interact with consistent selves.

So our justification of immanence as the root of choices stems both from the fact they are the ontic presence of alternative futures *and* that they are selected between by selves in a manner that is amenable to the attribution of responsibility.

We now turn to explain why it is essential that the exercise of free will arises from an interaction between a consistent self and an immanent relatum and not from interactions between the self and indeterministic, non-immanent, relata. So our question comes down to asking: how it is that a consistent self encountering an immanent relatum gives us warrant to attribute responsibility to the self for the outcome whereas the interaction between a self and an indeterministic relatum does not?

Consider first the scenario where a consistent self conjoins with an indeterministic relatum to jointly cause some consequent. As one of the relata has no determinate consequent, or even a delineated set of possible consequents in contrast to the immanent case, then the resulting caused relatum is undetermined by the combination. That is to say that the presence of the self may effect some aspects of the consequent, as after all it has causal input, but the complete result is down to serendipity alone: purely indeterministic aspects must always trump deterministic aspects. In such a situation can we legitimately place responsibility for the outcome onto the self? The answer must be that we cannot as the presence of the self does not lead to a unique outcome and does not give warrant for the attribution of responsibility. Informally, we might say that were we able to repeat the scenario then alternative outcomes could ensue and so the actual outcome, as we explained in section 5.10 above, cannot be the responsibility of the self. What we are claiming here is that having causal input into some outcome is not itself sufficient to give warrant to possessing responsibility for that outcome. If we were to allow any causal input to be sufficient for responsibility then we would have a free-for-all where every billiard ball, every mote of dust, every star could be said to be responsible for any relatum to which they causally contribute.

Finally, we turn to explain why we require the encounter between the self and the immanent relatum to arise serendipitously. This also explains why our mechanism can only take place within a universe that has indeterministic aspects to it - specifically indeterministic immanent relata. We start our explanation by setting up a counterfactual construal of our mechanism but one that does not involve deliberation: for our scenario, we might say that were we to repeat the scenario again and again then on each occasion the self would evolve according to its nature and conjointly cause a determinate outcome from the set of choices presented by the immanent relatum. If the immanent relatum had not encountered the self then it may have evolved differently. The reason it evolved the way it did, from the specific set of choices held within it, is that it interacted with a self and that interaction made the result an inevitability. In a certain sense we have moved the way we encapsulate free will in counterfactual terms away from making claims about the self, as we explained above, and instead shifted the counterfactual claim to the choices that are presented to the self.

For an immanent relatum A with immanent possibilia B and C encountering a self S, we can claim:

> **[FACTa]**: Relatum A on encountering S necessarily-caused B.

But, referencing the discussion in the previous chapter, it is true to say:

> **[SUPPa]**: If relatum A had not encountered S, it could have caused C.

What we claim is that the truth of **[SUPPa]** is in fact essential to our mechanism and is what makes the self *solely* responsible for the outcome of its choice.

Recall in section 4.6.1 we explained that immanent relata could arise in deterministic settings. Above we have ruled out responsibility arising from the indeterministic relata and instead said that responsibility for the selection from choices derives in part from ontic necessity. What this amounts to is the claim that the self is the external part of the universe that makes the immanent relatum it encounters into a deterministic one. That is, our mechanism itself relies on the form of deterministic immanence we discussed in section 4.6.1. We claim it is responsible for the choice because it made the selection at the *very same moment* the choices came into being - there was no time delay involved in our process. That is, the choice the self made was synchronous with the choices coming into being and it was the ontic necessity of the subsequent evolution that has given us partial justification for attributing responsibility to the self for its selection. Now, if the immanent relatum had encountered the self by necessity, i.e. if the encounter were made ontically necessary by earlier relata, then we can no longer claim that the self was *solely* responsible for the outcome. Rather, the earlier universe and its deterministic machinations seem to be equally culpable. Indeed, here we can recall van Inwagen's criticism of free will and determinism:

> If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things [including our present acts] are not up to us. (van Inwagen, 1986, p v)

If we wish to make the self *solely* responsible for the outcome then we must ensure there are no other aspects of the universe that led to the choices being presented to the self via ontic necessity. And this means we must require **[SUPPa]** to be true.

Now for **[SUPPa]** to be true, we must invoke a further immanent relatum. This relatum must contain the possibility that the immanent relatum that gave us our choices could have occurred elsewhere rather than encountering the self. And this can only happen if we take it that truth of **[SUPPa]** is to be found in an earlier indeterministic immanent relatum.

It is this then that adds the final part of the jigsaw that justifies our claim to have provided a version of the exercise of free will that offers genuine choices to a self and then permits us clearly to have the self choose an outcome from these choices in a manner that warrants the self being solely responsible for the outcomes.

After setting out a brief summary of our position we shall turn to examine one final question: whether we can ever know if we have free will of the sort we advocate here.

5.12 Review of the ontic exercise of free will

Here we offer an informal brief summary of the essence of our mechanisms for the exercise of free will. We review the mechanism again in Chapter 7.

Recall our two criteria of the ontic operation of free will from Chapter 1:

> **1.** The universe presents genuine choices, genuine alternative futures, to its inhabitants;
>
> **2.** The inhabitants 'choose' between these choices in such a manner that they are clearly and solely responsible for their choices, their selections.

In Chapter 4, we argued that an ontic notion of choice can only derive from immanent relata - ones that contain an embryonic presence of future relata.

In Theorems 2 and 3 above we argued that no set of relata can self-pilot or pilot other relata towards one choice rather than the other.

Putting these together shows that if we, as denizens of a universe, are presented with choices as immanent relata we cannot then pilot them towards one outcome over another.

If we cannot pilot genuine choices towards one outcome over another then we cannot define responsibility using any notion of such piloting. Hence we dismissed any deliberation-based counterfactual unpacking of free will.

Instead we proposed that responsibility comes from acting in accordance with one's nature. And we took it that to act in accordance with one's nature is to cause in a manner that is fully ontically determined by one's current constituting relata.

So if we are to act in a manner that warrants the attribution of responsibility we must act deterministically. And if we are to select between choices we must do so in a manner where the outcome is determined by our nature.

So to act responsibly is to force an immanent relatum to evolve by necessity into one of its inhering possibilities. The responsibility for the choice is then the result of the ontic necessity that the self brought to the immanent relatum in forcing it to move one way rather than the other. And the uniqueness [our 'solely'] of the responsibility is to be found in wider indeterministic aspects of the universe.

If we reject immanent relata we loose choices.
If we reject deterministic selves we loose responsibility.

The result of this analysis is that choices cannot be derived from self-piloting or pure indeterminism; rather choices appear as immanent relata. And responsibility cannot be derived from self-piloting but rather derives from a mechanism that brings ontic necessity to bear on choices.

Immanence is our route to genuine choices in satisfaction of **1**

Consistent selves causing deterministically are our route to responsibility in satisfaction of **2.**

5.13 Can we know if we have free will?

We ask if we can know if we possess free will of our variety. We frame our answer via a return to our discussions in Chapter 2. There, we held that we only have epistemic access to the content of the states of the universe and thus could not tell if we inhabited a deterministic or indeterministic universe of any variety, including Humean. However, on the assumption that we inhabit a connexion-bound universe, our assertion rested in part on the division between prediction and determinism - that neither is a requirement for the other, as we explained in 2.9. What we did not explicitly consider is the case where indeterminism resides not merely within the sequence of states and their connexions but that it affects the types of relata that can arise in the universe - such as immanent relata.

Our concern is this: if we permit the possibility of immanent relata, and admit we inhabit a connexion-bound universe, can we then tell if we inhabit a deterministic or indeterministic universe?

We ask whether the presence of immanent relata makes our claim that we cannot differentiate between a deterministic and indeterministic universe invalid. Our original claim hinged on a clear difference between predictability and determinism. However, we cannot use that argument here.

Let us attempt to respond to this issue by taking a look at what it would mean for us to have epistemic access to immanent relata. Consider an immanent relatum, $\mathcal{I}(t_0)$ at time t in which inhere two possible outcomes, A and B. Now, in order for a self to have epistemic access to this immanent relata, we invoke equation (5.1) to give:

$$\mu(\mathcal{I}(t_0), \Xi(t_1)) = \mu(\mathcal{I}(t_0)) \tag{5.8}$$

There are three ways in which the immanent relatum can impinge on the observing self here: it can have the causal force of A, or it can have the causal force of B, or it can cause a further immanent relatum that is firmly part of the self and obeys (5.8). If it does have merely the causal force of A or of B then it is detected by the self as being A or being B: that is, it is not detected as an immanent relatum at all. So, in order for the self to have access to, and hence detect, an immanent relatum then part of the self must become an immanent relatum itself.

Does this allow us to assert that we live in an indeterministic universe. The answer is 'no'. The reason is simple: we have explained, in section 4.6.1, that immanent relata can exist in deterministic settings so they are not exclusive to indeterministic universes. So, if we can indeed detect immanent relata, we cannot make the move from this fact to assertions that we inhabit an indeterministic universe.

What implications does this have on our first question: can we tell if we have free will? The answer, again, is simple: even if we admit to living in a connexion-bound universe, we cannot tell if it is deterministic or indeterministic and so cannot tell if we might possess free will of our variety - one that requires both immanence and indeterminism.

5.14 The Libet experiments

As a final comment it is useful to say something briefly, and slightly informally, about how our thoughts might play a part in the exercise of free will.

We start by quoting Libet at length:

> How the brain deals with voluntary acts is an issue of fundamental importance to the role of conscious will and, beyond that, to the question of free will. It has been commonly assumed that in a voluntary act, the conscious will to act would appear before or at the start of the brain activities that lead to the act. If that were true, the voluntary act would be initiated and specified by the conscious mind. But, what if that were not the case? Is it possible that the specific brain activities leading to a voluntary act begin before the conscious will to act, in other words, before the person is aware that he intends to act? This possibility has arisen partly from our evidence that sensory awareness is delayed by a substantial time period of brain activities. If the internally generated awareness of the will or intention to act also is delayed by a required period of activities lasting up to about 500 msec, it seems possible that the brain's activities that initiate a willed act begin well before the conscious will to act has been adequately developed.
>
> We were able to examine this issue experimentally. What we found, in short, was that the brain exhibited an initiating process, beginning 550 msec before the freely voluntary act; but awareness of the conscious will to perform the act appeared only 150-200 msec before the act. The voluntary process is therefore initiated unconsciously, some 400 msec before the subject becomes aware of her will or intention to perform the act. (Libet, 2005, pp. 123-124)

We are interested in noting that Libet suggests that the conscious awareness of a decision may well occur after the initiating event has occurred and so play no fundamental part in the making of the initial decision.[11] If we do indeed have free will in our sense, then the conscious awareness we have of possible choices and the making of these choices might not be ontically reflective of what is actually happening: whilst I might have the impression that I have consciously decided to make one choice rather than another that impression must be, to some extent, an illusion in that the decision to do A or to do B was taken prior to the conscious awareness's becoming involved. This is entirely in keeping with what we claim here. We are not claiming that the conscious self is a participant in the initial interaction with an immanent relatum. If anything, the conscious process is perhaps, in part, a reflection of the machinations of the self that has already made the initial selection. Perhaps the conscious processes involved are an integral part of the causal processes that lead from initial interaction between a self and an immanent relatum to the actions - the moving of arms, the issuing of speech etc - that ensue as a result. So we can state then that, whilst Libet's experiments may suggest that the conscious mind comes into play only after decisions and choices are made, it does not then follow that the thought processes themselves are not an integral part of the self or are not an integral part of the process that passes from choice to action.

[11]For various accounts and perspectives on this, see Libet (1996), Libet et al. (2000), Libet (2005) and Libet (2006).

Chapter 6

Supervenience

6.1 Introduction

In this chapter we look at supervenience.

Recall in Chapter 1 we set ourselves three broad aims. One was to set up mathematical models of some concepts that tend to feature in discussions on free will. Another was to use these models to construct a version of free will. And the third was to lay foundations for future work. This chapter contributes, in varying proportions, to all three aims.

The chapter starts with a brief but rigorous treatment of supervenience. In developing our model of supervenience here we supply sufficient detail and motivation to deal with most, if not all, versions of supervenience appearing in the literature. The approach we take is deliberately general but enough examples and hints are given to accommodate multifarious forms of supervenience and deal with technical theorems and ideas in an efficient and rigorous manner. In other words, we set up enough mathematical structure to act as a strong foundation for further study. This aspect of the chapter fulfils our stated aim of constructing mathematical models of areas related, in broad terms, to free will and to our aim of setting out foundations for future study.

Following our model of supervenience, we move to link it with the work we have outlined in previous chapters. We link together the causal model we set up in Chapter 3 with the supervenience model we set up in this chapter. We do this for two reasons: one reason, which only comes to fruition in the very last section of the chapter, is to demonstrate that considerations of supervenience do not clash against the causal arguments we have set out in this study. In this last section we conclude that, in most cases, supervenience does not provide us with any worries. Much of the technical work in the second half of the chapter - where we merge causal concerns and supervenience - is set up with this first aim in mind. The other reason we link causality and supervenience is that by doing so we can clarify in what way supervenience might supply a synchronous ontic relationship which is not merely a shadow of some causal relationship. This serves one of our stated aims: to set up foundations for further work. Indeed, we shall show that there is a potential version of supervenience that does not derive from causal concerns - we call it 'pure supervenience'- and we make some tentative suggestion that it might be linked to some notion of emergence. We do not, however, pursue emergence in detail.

6.2 Defining supervenience

6.2.1 The essence of supervenience

Supervenience is a philosophical idea that has had much prominence over the years, particularly in aspects of the mind-body debate.[1] In addition supervenience is extensively developed by Kim,[2] and has found itself an important part of the discussion concerning the mind and the world. As Kim (1993) says,

> ... think of supervenience as belonging in that class of relations, including causation, that have philosophical importance because they represent ways in which objects, properties, facts, events, and the like enter into dependency relations with one another, creating a system of interconnections that give structure to the world and our experience of it. (Kim, 1993, p 54)

A good and simple way of appreciating the idea of supervenience is to start with Davidson (2001). He says,

> Mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be **[A]** two events alike in all physical respects but differing in some mental respect, or **[B]** that an object cannot alter in some mental respect without altering in some physical respect.[3] (Davidson, 2001, p. 214)

Davidson's definition can be used as a launch to help build up an intuitive picture of supervenience. We can take it here, at least as a convenience for exposition,[4] that his claim is that there is a

[1]For a very useful summary of the subject see McLaughlin & Bennett (2011) and for a survey of the subject see Savellos & Yalcin (1995).

[2]See Kim (1993). Much of the first half of this chapter draws on his definitions and approach. Also, important discussions are found in Teller (1984), Paull & Sider (1992), Horgan (1993), Chalmers (1996, pp. 32-42), Stalnaker (1996) and Butterfield (2011).

[3]**[A]** and **[B]** are added here.

[4]Davidson's unpacking is much more subtle than we imply here. We simplify considerably as we are using his definitions solely as a segue into a general notion of supervenience.

relationship between a person's mental states,[5] on the one hand, and the contemporaneous physical states of their brain on the other.[6] We shall take the relationship can be taken in two, equivalent, ways:[7] one derived from section **[A]** and one from section **[B]** of the quote, skirting over the variation in application to events in **[A]** and objects in **[B]**.

Consider Janet and John and their brains.[8] Section **[A]** says that if we were to have some method of examining their brains down to the minutest detail at some time and that in doing so we discover the brains were identical in their structure, down to the relative placement of every molecule, then we could say that they had identical mental states at that moment. So, for instance, both Janet and John would then be harbouring identical desires to eat an ice cream placed in front of them.

For section **[B]** we have to consider a slightly different approach. We ask both Janet and John what they are thinking at some moment,[9] so that Janet talks of ice-cream and John of éclairs. At this moment we can be sure that their mental states are different, then accepting section **[B]** amounts to accepting the fact that the physical make up of their brains must also be different.

A third case serves to illustrate the restricted scope of the supervenience relation. We now ask Janet and John a further question and discover that at the moment of our asking they are entertaining absolutely identical thoughts of eating a plate of cheesecake. They have identical mental states but the supervenience relation does *not* allow us to say that they must then have identical physical brain states. Instead, the question as to their brain states remains an open question.

So here we have a condensed essence of the idea that the mental-M supervenes on the physical-P: section **[A]** says that P-identity requires M-identity; and **[B]** says that we cannot have M-differences without accompanying P-differences. And importantly, supervenience does *not* say that M-identity requires P-identity.

We have deliberately left some matters vague in this section and oversimplified matters for the purpose of exposition. For instance, we have taken it that our objects are obviously associated so that 'John', 'John's brain' and 'John's thoughts' are all associated together. And similarly for Jane etc.

---

[5] Davidson restricts his mental states to dispositional attitudes. See Davidson (2001, pp. 209-210).

[6] We shall take it throughout this chapter that any association between physical and mental states is taken at specified times. This can possibly be taken to contradict with empirical evidence, for instance Libet et al. (2000), Libet (2006) and Klemm (2010), but this does not affect the central thrust of the arguments presented. We examined an aspect of Libet's ideas and experiments in section 5.14.

[7] We ignore the possibility that by 'alter' Davidson intends to inject some time based aspects into his idea of supervenience. The methods we develop in section 6.2.7 can be used to modify our static model of supervenience into one that can accommodate alterations in time. As such we couch our model in terms that take **[B]** as the contrapositive of **[A]**.

[8] We take the stance that we can talk of mental states supervening on brain states as a convenient expository tool here. We do *not* claim that this is what Davidson holds to or that it is the only way of cashing out supervenience relationships between the mental and the physical.

[9] The use of 'thinking' is a convenient oversimplification here.

6.2.2 Formal supervenience

In this and following sections, we shall build on the intuitive ideas outlined above and set up a formal mathematical articulation of a basic supervenience relation. This will be developed and refined in subsequent sections. We start with a set of 'objects' which are to be part of the focus of the supervening relation and to set its scope. The set of objects may be a set of brains of a group of people at a specified moment, or the brain of an individual at various times, or lumps of matter, or cats or even just a more abstract collection, such as a set of properties, events or some such. The point is that not too much baggage should be attached to the word 'object' at this stage. Initially the individual objects will each be taken to be dependent upon a region of space,[10] Δx, and to be specified at a particular time, t, and at a particular world w. It may be the case that when investigating concrete cases of object sets that these dependencies fall away, or perhaps become issues of contention, but that is not important yet.

We define a set of objects:

$$\mathbb{O}_i = \{O_j(\Delta x, t, w_k); j \in \mathcal{I}\} \tag{6.1}$$

with $\mathcal{I}$ the indexing set and t and k fixed. As supervenience is usually taken to be a relationship of similarity and difference amongst sets of properties [of objects], we shall need to consider sets of properties that can be possessed by these objects. Properties will *not* be indexed in any way by worlds, time or space as they will be taken as applicable in any world, time or place they are specified:[11]

$$\Pi_l = \{P_m; m \in \mathcal{I}\} \tag{6.2}$$

Properties are taken to be maps from the set of objects to the numbers 0 and 1. That is to say, we take our objects as being the ontic structures that define how properties are grouped together. In simple terms an object is mapped to 1 by a property if it possesses a property associated with the map and it is mapped to 0 if it does not possess a property associated with the map. We can write this formally:

$$P_m(O_j) \in \{0, 1\} \tag{6.3}$$

[10] Space plays almost no role in our version of supervenience. We include it here for completeness.

[11] For the moment, we use subscripts to differentiate between properties. In section 6.3.5 this will be used to differentiate between general properties, or property-types, and specific examples of properties, or token instances.

We then set up a definition of object equality:

DEFINITION 6.1: Identical with respect to properties

Two objects will be said to be **equal [or, identical] with respect to some property** if they both give the same value when they are arguments of that property.

For instance, the earth and a ping-pong ball are equal or identical with respect to the property of being-spherical. Each property acting on a set of objects can be used to divide the set into two disjoint subsets,[12] that are collectively exhaustive,[13] on the assumption there is never any doubt as to whether an object possesses a well-defined property.[14] We can set up equivalence classes with respect to some property P_i using:

$$O_j \sim_i O_k \tag{6.4}$$

Where $\sim_i$ is taken to say that two objects are in the same equivalence class if they are identical with respect to property P_i. So a set of properties acting on a set of objects will produce identity groupings amongst the objects. We shall tighten up this in a short while.

As an example, consider a set of objects comprising spheres and cubes, some of which are painted red and some blue. The four properties spherical, cubic, red and blue serve to divide up these objects into various groupings. So that if we were to consider only the spherical property, we would *not* be able to distinguish between red spheres and blue spheres. If we were to consider both the property being-spherical and the property being-red then we could distinguish red spheres. Of course, we cannot use the sphere property to distinguish cubes from spheres. Knowing that an object does not possess the sphere property is not sufficient for us to say it must therefore be a cube. For an object to be a cube it must possesses the cube property. That is not-being-spherical tells us nothing else about the object at all: it does not mean it must be cubic any more than it means it must be red or blue. Of course, we may discover that the property of not-being-spherical is the same as the property of being-cubic in this situation but that is not to say much. This particular situation does not tell us much else about the spheres and cubes. It might be the case that some spheres are the size of the earth and others are the size of ping-pong balls. Whilst we inhabit this particular set up these sorts of distinction are beyond us as we do not have access to the properties we need to make other sorts of distinctions between objects.

[12]Two sets are disjoint if they have no elements in common.

[13]A set of subsets is exhaustive if each element in question belongs to at least one of the subsets.

[14]We say this as there are complications from some interpretations of quantum theory. In some interpretations, the possession of a property is not guaranteed by the object alone and is dependent of other aspects of the set up. For a brief discussion, see Isham (1995).

We can now define what it means to say two objects are **identical** or **not identical** [or **indiscernible** or **indistinguishable** and so forth], when we come to talk of supervenience. To tighten up this definition we shall define the 'distance' between objects, from a specified object set $\mathbb{O}_i$, with respect to a specified set of properties Π_l, as follows:

$$\mathfrak{d}_{i;l}(O_j, O_k) \equiv \; max\|P_m(O_j) - P_m(O_k)\| \; \forall P_m \in \Pi_l \tag{6.5}$$

Where the subscripts on $\mathfrak{d}_{i;l}$ correspond to the label for the set of objects, the 'i', and the set of properties, the 'l'. And then we can divide up the objects into equivalence classes according to

$$O_j \sim_{i;l} O_k \Leftrightarrow \mathfrak{d}_{i;l}(O_j, O_k) = 0 \tag{6.6}$$

Equation (6.6) amounts to saying that two objects are to be considered the same with respect to our property set if they are indistinguishable with respect to the property set. That is two objects are indistinguishable if there is no property in the set that one possess that the other does not. So (6.6) divides up our objects into groupings of indistinguishable objects. In our earlier example it would divide the cubes and spheres into four groupings, namely red-spheres, blue-spheres, red-cubes and blue-cubes. This does not mean grouped objects are to then be the same *simpliciter* as there may be further properties that we have not included in our property set which could be used to distinguish them, as was the case with our example of the earth and a ping-pong ball.

A set of objects divided up in this way will be called a **pre-topology** and denoted as:

$$\boldsymbol{T}_{i;l} = \{\mathbb{O}_i, \Pi_l\} \tag{6.7}$$

What this amounts to is that two objects in a pre-topology can be distinguished *with respect to the specified property set* if there is some property that one possesses that the other does not.

Before proceeding it is worth re-emphasising that a pre-topology requires both a set of objects *and* a set of properties for its construction. However it is not the case that if we change the property set of these a new pre-topology results as the pre-topology is only the set of groupings of the specified objects by the specified properties and it is possible for varying property sets to produce the very same groupings. For instance, if we were to use the property set consisting of some property F or use the property set consisting of not-F the resulting divisions of the object collections would be indistinguishable and hence the pre-topologies would be too. So we shall introduce a further definition that we may call upon when distinction between property sets is needed. We shall call a **specification** a triple consisting of a set of objects, a set of properties and the induced pre-topology.

This will be denoted as:

$$S_{i;j} \equiv \{\mathbb{O}_i; \Pi_j; \boldsymbol{T}_{i;l}\} \tag{6.8}$$

This notation then allows us to make finer distinctions than the pre-topology definition alone. For instance, as we mentioned, the property F and the property not-F each induce the same pre-topology on a given object set so using pre-topologies alone does not allow us to take account of the differences between F and not-F. The specification, however, does. Two specifications are equal if they have both the same set of objects and the same set of properties. Most of the time, this sort of distinction will be clear and talk of pre-topologies alone will suffice so we shall have no explicit need for specifications in what follows.

When we come to define supervenience below some distinctions will be important. For instance, we cannot usually talk of one set of properties supervening on another set of properties without first articulating what objects are to be involved and this will have implications later.

We are now in a position to start shaping a definition of what it is for one set of properties to supervene upon another. Our initial definition is to be thought of as a launch for more precise definitions later and it will skirt round a number of issues which will reappear for clarification.

We shall start with two pre-topologies which have their objects in common but not their properties:[15] $\boldsymbol{T}_{i;a} = \{\mathbb{O}_i, \Pi_a\}$ and $\boldsymbol{T}_{i;b} = \{\mathbb{O}_i, \Pi_b\}$. We shall not concern ourselves here with the explicit indexing of objects by time, space or worlds and just take it that we have a static set of objects at a given world at a specified time. This is very much like our example in section 6.2.1 where we took our objects to be, albeit vaguely, Janet and John and the two property sets to be, once again rather vaguely, mental states and physical brain states.

We shall take our definition of supervenience as coming from **[A]** in the previous section although, as we shall show, this is equivalent to **[B]**:

DEFINITION 6.2: Supervenience of property sets

We say that one set of properties, Π_a, **supervenes** on another set of properties, Π_b, with respect to a set of objects, $\mathbb{O}_i$, when the corresponding objects in the pre-topologies obey the following condition:

$$O_j \sim_{i;b} O_k \Rightarrow O_j \sim_{i;a} O_k \tag{6.9}$$

[15] We should really talk of the pre-topologies as part of two specification which have differing property sets. However referring to the pre-topologies directly, although slightly casual, should not cause undue confusion. In future, we shall not feel the need to comment on this shorthand way of talking unless the distinction needs to be emphasized.

For ease we shall usually talk of $T_{i;a}$ supervening upon $T_{i;b}$ and we shall write this as:

$$\boldsymbol{T}_{i;a} \looparrowright \boldsymbol{T}_{i;b} \tag{6.10}$$

Where (6.9) amounts to saying the same thing: that supervenience says that if any two given objects are indistinguishable with respect to property set Π_b then they are indistinguishable with respect to property set Π_a.

But, of course, we must be careful with such talk as the supervenience relationship is between sets of properties with respect to a given set of objects. Our notation should be clear: when we talk of pre-topologies or, later, topologies alone we mean the objects considered as grouped but when we talk of pre-topologies or topologies supervening then we are strictly talking about the relationship between property sets with respect to object sets and defined using the corresponding topologies. We justify the blurring of the notation here by the fact that what is meant should be clear in the corresponding context and it avoids the constant use of clarifying phrases.

Recall our rough motivating definition from Davidson above:

> Mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect. (Davidson, 2001, p. 214)

And recall, (6.9) amounts to saying the same thing, that is that supervenience says that if any two given objects are indistinguishable with respect to property set Π_b then they are indistinguishable with respect to property set Π_a. We see this is equivalent to the **[B]** style of supervenience from above by taking the contrapositive:

$$\neg(O_j \sim_{i;a} O_k) \Rightarrow \neg(O_j \sim_{i;b} O_k) \tag{6.11}$$

Where we take it that:

$$\neg(O_j \sim_{i;a} O_k) \Leftrightarrow \eth_{i;a}(O_j, O_k) = 1 \tag{6.12}$$

(6.12) says that two objects are distinguishable with respect to property set Π_a if there is at least one property in the set that one object possesses and the other does not possess.

The supervenience relation is both reflexive and transitive:

$$\boldsymbol{T}_{i;a} \looparrowright \boldsymbol{T}_{i;a} \tag{6.13}$$

$$(\boldsymbol{T}_{i;a} \looparrowright \boldsymbol{T}_{i;b}) \wedge (\boldsymbol{T}_{i;b} \looparrowright \boldsymbol{T}_{i;c}) \Rightarrow (\boldsymbol{T}_{i;a} \looparrowright \boldsymbol{T}_{i;c}) \tag{6.14}$$

And transitivity is easy to prove from the definition of supervenience above:

$$\begin{aligned}(O_j \sim_{i;b} O_k \Rightarrow O_j \sim_{i;a} O_k) \wedge (O_j \sim_{i;c} O_k \Rightarrow O_j \sim_{i;b} O_k)&\\ \Rightarrow (O_j \sim_{i;c} O_k \Rightarrow O_j \sim_{i;a} O_k)&\end{aligned}$$

But supervenience is not necessarily symmetric.

6.2.3 The completion of a property set

We now turn to the consequences of augmenting the property set in a given specification. The long-term motivation behind this is that it facilitates more complex trans-world definitions of supervenience which can then be used to make more subtle philosophical claims about property sets. We shall meet these trans-world definitions in section 6.2.6 below. For the moment, the main motivation for augmenting property sets will come from a desire to introduce the idea of *entailment* into supervenience. Later, entailment also plays a part in widening the scope of the notion of supervenience. Essentially, the augmentation is achieved by assuming a property set is closed under a certain set of operations which we shall define below.

We start with a definition of entailment:

> **DEFINITION 6.3: Property entailment**
>
> A property P_j is **entailed** by a property P_k within a pre-topology $\boldsymbol{T}_{i;a}$ or across pre-topologies with identical object sets if, for all objects in $\mathbb{O}_i$, it is the case that $P_k(O_m) = 1 \Rightarrow P_j(O_m) = 1$.

What this amounts to is that entailment of properties requires the objects considered as identical with respect to P_k must also be identical with respect to P_j. We can write this suggestively as $P_k \subseteq P_j$.

Now we can ask what is the relation between supervenience and entailment; that is, does supervenience necessarily involve entailment. The simple answer is 'no' and a straightforward example will suffice to show this. Consider the two pre-topologies constructed each from the same set of objects but one having a single property, P, and the other its negation, $\neg P$. It is easy to see that each pre-topology supervenes, or more correctly each set of properties supervenes, on the other with respect to our specified object set, but it is not the case that P entails $\neg P$ or even vice-versa. This example also emphasizes our earlier point too, that entailment, like supervenience, is a relationship between property sets with respect to a given object set and not a relationship between sets of objects - recalling that the objects are necessary to define the scope of the supervenience relation. That is, although it is the case that each of our pre-topologies contains sets that are [improper] subsets of objects in the other pre-topology that alone does not mean that entailment ensues.

Perhaps we can ask if given a supervenience relation between pre-topologies, we can then augment one or both of the property sets so that some entailment relations are guaranteed between the property sets.

In order to see how entailment can be guaranteed in some form given a supervenience relation we turn to the [Boolean] completion of property sets as suggested by Kim (1993). We start by creating an augmented property set by allowing for all possible negations, conjunction and disjunction of properties.[16] So we have the completion of a property set, defined as:[17]

$$\overline{\Pi}_l = \{P_m \ldots ; \neg P_m \ldots ; P_m \cap \neg P_n \cap \ldots ; P_m \cup \neg P_n \cup \ldots etc\} \tag{6.15}$$

Together with:

$$(P_m \cap P_n)(O_i) = Min\,\{P_m(O_i), P_n(O_i)\} \tag{6.16}$$

$$(P_m \cup P_n)(O_i) = Max\,\{P_m(O_i), P_n(O_i)\} \tag{6.17}$$

$$\neg P_n(O_i) = 1 - P_n(O_i) \tag{6.18}$$

Where these three equations, (6.16) to (6.18), are the extensions of the rule we outlined in (6.3) and **DEFINITION 6:1** above. Recall, this definition said:

[16] Kim allows infinite conjunctions and infinite disjunctions for technical reasons and we shall allow countably infinite combinations, even though there may well be problems with this. For instance, when considering the property of being a closed set of the real line and allowing infinite intersections. We shall ignore such difficulties here.

[17] Some accounts use logic notation for the combination of properties: using $\wedge$ for 'and' and $\vee$ for 'or' and $\neg$ for 'not'. Here we use mostly set theoretic notation - with the exception '$\neg$' for 'not' - in place of this because we shall also be using use logic notation conventionally below and we want to avoid confusion. We take it that our notation is directly translatable as follows: $\wedge = \cap$ and $\vee = \cup$.

DEFINITION 6.1: Identical with respect to properties

Two objects will be said to be **equal [or, identical] with respect to some property** if they both give the same value when they are arguments of that property.

We decided there to set the outcome of a property acting, as a mapping, on an object to be 1 if the object possesses the property and 0 otherwise. So, (6.16) says that (O_i) gives the value 1 if and only if it possesses both P_m *and* P_n; similarly, (6.17) says (O_i) gives the value 1 if and only if it possesses either P_m or P_n or both; and, (6.18) says that (O_i) gives the value 1 if and only if it does *not* possess property P_n.

This process of combining properties is not without philosophical controversy. For instance, it is not clear that the conjunction of two properties is itself legitimately a property and nor does it always seem sensible to allow the negation of a property to also be a property:[18] a giraffe is not an asteroid but so what? It is worth mentioning that this may allow for the distinction between forms of logical supervenience, where the combination of properties is to be allowed, and forms of nomological supervenience where property combination may well not be allowed.[19]

The resulting property set (6.15) is then closed under the operations of negation, conjunction and disjunction. That is to say that we cannot form further properties from the completed set using these operations. Also, as every object either possesses a property or its negation then the completed property set must contain both the empty-property, P_ϕ, and the universal-property, P_Π:

$$P_\phi(O_j) = 0\,;\; P_\Pi(O_j) = 1\;\forall O_j \tag{6.19}$$

We shall call the new pre-topology constructed from the completed set of properties a **topology** and any pre-topology induced by a property set that is closed under the Boolean set of operations will be called a topology to distinguish it from the related pre-topologies. There is a subtle, and not entirely obvious, point to make here: that is that the groupings of objects in a pre-topology and its closure are exactly the same, the difference is only that the topology comes from an augmented set of properties. So that the distinction we make here between a pre-topology and a topology by placing a bar across the symbol, actually only has relevance when we wish to talk about the prop-

[18]We make some further comments on this below in section 6.3.5.

[19]There is an additional problem that is apposite to mention here: the use of properties in some interpretative versions of quantum theory. There it is often maintained that the objects under consideration, such as electrons, photons and so forth, do not possess properties prior to measurement, that is the properties are 'brought into being' by the act of measurement; further, the Heisenberg uncertainty principle is often taken to say there is an ontological restriction on what definite properties quantum objects can possess at any one time. So that an electron that possesses a definitive position in space then cannot possess any specified momentum. We shall ignore these issues for the moment and take it that property combinations are allowable under restricted circumstances and not make much of what these circumstances may be.

erty sets that induce the object groupings and it does not affect the object groupings themselves. We can write:

$$\overline{\boldsymbol{T}}_{i;a} = \{\mathbb{O}_i, \overline{\Pi}_a\} \tag{6.20}$$

But we note, considered purely as division of objects into groups, it is the case that:

$$\overline{\boldsymbol{T}}_{i;a} = \boldsymbol{T}_{i;a} \tag{6.21}$$

This shows us that the Boolean completion of property sets does not directly affect the topologies concerned and so allows us to make a number of claims about the supervenience relationship between property sets. For instance, all of the following are valid together, even though each involves different sets of properties:

$$\boldsymbol{T}_{i;a} \looparrowright \boldsymbol{T}_{i;b} \tag{6.22}$$

$$\overline{\boldsymbol{T}}_{i;a} \looparrowright \boldsymbol{T}_{i;b} \tag{6.23}$$

$$\boldsymbol{T}_{i;a} \looparrowright \overline{\boldsymbol{T}}_{i;b} \tag{6.24}$$

$$\overline{\boldsymbol{T}}_{i;a} \looparrowright \overline{\boldsymbol{T}}_{i;b} \tag{6.25}$$

That is, we have:

$$\boldsymbol{T}_{i;a} \looparrowright \boldsymbol{T}_{i;b} \Leftrightarrow \overline{\boldsymbol{T}}_{i;a} \looparrowright \boldsymbol{T}_{i;b} \Leftrightarrow \boldsymbol{T}_{i;a} \looparrowright \overline{\boldsymbol{T}}_{i;b} \Leftrightarrow \overline{\boldsymbol{T}}_{i;a} \looparrowright \overline{\boldsymbol{T}}_{i;b} \tag{6.26}$$

In other words, we are given a property set that supervenes on another with respect to some object set, as in (6.22). We then create two further sets of properties out of the old ones by making them Boolean complete. We find there are supervenience relationships between the new and old sets - such as in (6.23) to (6.25) - inherited from the original supervenience relationship as expressed in (6.22).

The completion does allow us to make additional claims concerning entailment. Indeed, we can now answer the questions we posed earlier. If we take a supervenience relationship $\boldsymbol{T}_{i;a} \hookrightarrow \boldsymbol{T}_{i;b}$ then as we have seen there is no guarantee that any property in Π_b entails a property from Π_a. However, if we take the closure of $\boldsymbol{T}_{i;b}$, which will still preserve the supervenience relation, then there is an additional guarantee that some property in Π_b will entail a property in Π_a. And, as mentioned, the transition from a pre-topology to a topology preserves the supervenience relation as it does not change the groupings of the objects. This is because the Boolean completion of the property set does not add or detract from the distinctions amongst objects that the original property set instantiates. That is to say that both the original set of properties and their Boolean completion, divide the set of objects up into identical sets of indiscernibles. The only additional distinction is that the Boolean completion allows each set of indiscernibles in the object set to be associated with a single property whereas this is not necessarily the case for the original property set. The literature on supervenience usually requires that it is a relationship between Boolean completed sets.[20] Here we prefer to add the Boolean completion as an extra condition.

6.2.4 Object sets, time and other approaches to supervenience

Before we move to look at some examples from the literature and examine how they fit with the structures we have developed above, we need to make some comments. First, in our model we took supervenience to be defined with respect to specified object sets at specific times. That is our notion of supervenience was time specific and reliant on what object sets we decided to use when investigating the relationships between the property sets. Further, we held that properties were independent of time and world, and we did not require our property sets to be Boolean completed. In contrast, the literature does not do this. First, it appears that the literature takes the object set to be the total of all objects extant in a universe - we put the time aspect of this to one side for a moment. Further, it is usually the case that the property set, whilst applying - as ours do - equally at any time or world, is taken to be Boolean completed. There is, however, a third issue with standard approaches to supervenience that might prove to mean that some, and perhaps most, of the definitions in the literature are ambiguous.[21] When supervenience is unpacked it appears there is generally no mention of the time at which the supervenience relationship is to apply. That is, it may well be the case that the object sets are taken to be *all objects* that exist at *all times* within a universe rather than our, somewhat more restricted set; recall, all our definitions were time specific.

[20]For instance, see Kim (1993, p. 58).
[21]We leave a detailed analysis of this for another time.

Why might this be a problem for some approaches to supervenience? It is a problem as it is not clear how to deal with what it is to be a persisting object - i.e. notions of trans-temporal sameness for objects; should we use a property based definition or some other criteria and what assumptions do these criteria bring to the mix? And also it is not clear how to deal with the way objects change their properties moment to moment. For instance a billiard ball may have a velocity of 4m/s at one moment and be stationary later. In such a cases, what are we to do with the variation in the velocity of the billiard ball if we are also to treat it as a single object? Equally, when dealing with trans-world, and intra-world, maps of objects, how are we to understand what time restrictions, if any, we place on our maps? In essence, it would appear that many approaches to supervenience lack a robust and clear delineation of the object set from which the supervenience relations gain their traction and this might lead to some ambiguity.

We remedy this difficulty by making some assumption that we shall keep to for the next few sections. We shall assume that the literature definitions apply with respect to an object set only at a specific time and this object set constitutes the complete set of extant objects in a universe at that specific time. That is we assume the literature definitions are in line with our time-specific definitions.

When we come to examine supervenience and causality in section 6.3 we relax this condition. We discuss how we achieve this relaxation in section 6.3.1.

6.2.5 Some types of supervenience

So far we have defined a basic notion of the supervenience relation with respect to a set of objects without worrying too much about the parameters we used to label objects, namely the spatial region occupied by the object, the specific time at which the property operates and the world in which the property operates. And throughout we have made the assumption that properties are both definite maps - that is, every object either does or does not possess a given property - and that properties have a trans-world presence so we can talk freely about properties without having to talk of times, or places or even of worlds. The variety of definitions of supervenience that appear in the central literature take it that the property sets are Boolean complete,[22] but we shall not take that as a required condition of the property sets. With these preliminary provisos, we now explore briefly some possible varieties of supervenience. We do this to provide general hints as to directions in which we could develop the ideas in a more extensive treatment than we provide here. The variety of supervenience comes primarily by exploiting the indexing of objects by worlds but continuing to assume properties are independent of both time and spatial location. We shall take it that when

[22]See Kim (1993, p.58).

we talk of supervenience amongst worlds that the set of worlds, W, that is under consideration is specified either explicitly or implicitly.

Here we look briefly at three types of supervenience: nomological, logical and 'metaphysical'. We mention nomological and logical to show how the supervenience relationship itself can have its roots in different features: the nomological in the laws of nature; the logical in the logical relationships. We also introduce a 'metaphysical' type of supervenience - which is not the same as that in the literature, where metaphysical is often taken as the same as logical - to demonstrate the need to move to trans-world definitions of supervenience, something we set out in section 6.2.6 below.

First, nomological supervenience. We take any nomic restrictions to be world specific and so nomological supervenience can only be defined *within* a specified world. We say that $\boldsymbol{T}_{i;a}$ nomologically supervenes on $\boldsymbol{T}_{i;b}$ if there is some world $w_m \in W$ such that:

$$O_j(w_m) \sim_{i;b} O_k(w_m) \Rightarrow O_j(w_m) \sim_{i;a} O_k(w_m) \tag{6.27}$$

And we write this as

$$\boldsymbol{T}_{i;a} \looparrowleft_N \boldsymbol{T}_{i;b} \tag{6.28}$$

We might want to make the nomological restriction apply to at least our world, that is the actual world we are taken to inhabit, and it should be clear from context whether the nomological supervenience refers to our world, amongst others perhaps, or some other world. Nomological supervenience is an attempt to encapsulate certain physical laws of the world we inhabit, or even a world we could imagine with slightly varying physical laws to our own, and expresses the relationship between various physical features. For instance, it may be the case that the mental supervenes on the physical nomologically in that we might take it that whilst it is true in our world, there may well be worlds where the supervenience relation does not apply. The classic example of the later is a world where there are zombies - who are taken to have no mental content but to behave in a manner that they cannot be distinguished from individuals who do posses mental content.[23]

Second, logical supervenience. Logical supervenience involves logical relations. So, for instance, X logically supervenes on Y if Y-ness entails X-ness. For instance, $A \cup B$ supervenes on the property set A and B - thus being-blue-or-green supervenes on the set of properties consisting of

[23] For a discussion, see for instance Chalmers (1996).

being-blue and being-green.

We must allow for logical completion of the supervening pre-topologies. So we ought to restrict the idea of logical supervenience to topologies. We denote logical supervenience by placing a subscript on the supervenience symbol, $\hookrightarrow_L$. We shall have no explicit use for logical supervenience in what follows.

Third, an ersatz metaphysical supervenience. Recall, this is to be used here mainly to illustrate the need to move to trans-world definitions of supervenience - something we do in the next section. Ideally metaphysical supervenience ought to have a stronger modal force than nomological supervenience, so we take it that metaphysical supervenience is valid in all the possible worlds that are under consideration but does not require any cross world supervenience. We say that $\boldsymbol{T}_{i;a}$ metaphysically supervenes on $\boldsymbol{T}_{i;b}$ if at all worlds $w_m \in W$:

$$O_j(w_m) \sim_{i;b} O_k(w_m) \Rightarrow O_j(w_m) \sim_{i;a} (O_k)(w_m) \tag{6.29}$$

And we write this as

$$\boldsymbol{T}_{i;a} \hookrightarrow_M \boldsymbol{T}_{i;b} \tag{6.30}$$

However, metaphysical supervenience, in this form, is less useful than it may initially appear. This is because the way in which the property sets divide up the objects *within* each world must maintain the supervenience relation between the property sets there but there is no guarantee that, in the case where we have some notion of trans-world identification of objects, that the objects possess the same properties from world to world. For instance, in our example of coloured spheres and cubes, we may take it that in one world the property of being spherical is possessed by spheres and being cubic is possessed by cubes, but in another world these are, in some sense, reversed, so that new-cubes are spherical and new-spheres are cubic. In this case, whilst the objects can take on different properties from world to world, the supervenience relation is untouched. It may be the case that this could be related with intrinsic properties, so that we require intrinsic properties of objects to be stable across worlds. Metaphysical supervenience, however, tells us nothing of great interest about properties and objects. This is because, as noted, it does not require there to be any preserved trans-world relationship between individual objects and individual properties. This suggests some manner of trans-world restrictions need to come into play if we are to set out more interesting forms of supervenience. Some more robust versions of trans-world supervenience are outlined in the sections 6.2.6 and 6.2.7.

In a more rigorous treatment of the ideas in this section a considerable number of distinctions could be made by extending the ideas of closure, trans-world object sets, and so forth. We leave this section as a hint of what could be achieved were we to develop the mathematical structures and definitions rigorously in those directions.

6.2.6 Weak and strong supervenience from the literature

There are a number of interlocking definitions that appear throughout the literature, and many of the important ones appear in a series of papers published by Kim over a period of twenty years or so.[24] In this, and the next, section some of Kim's definitions are set out and translated into our mathematical framework. Kim (1993) defines weak supervenience as follows:

> A weakly supervenes on B if and only if necessarily for any x and y if x and y share all properties in B then x and y share all properties in A - that is, indiscernibility with respect to B entails indiscernibility with respect to A. (Kim, 1993, p. 58)

This is defined using the Boolean completed sets of properties, although it stands as a definition without this stipulation in our structure, and it is taken to apply only within one world. Ignoring Boolean completion, we can translate this. We take the phrase, 'if x and y share all properties in B then x and y share all properties in A' to translate to:

$$x \sim_{i;B} y \Rightarrow x \sim_{i;A} y \tag{6.31}$$

which gives:

$$\boldsymbol{T}_{i;A} \looparrowleft \boldsymbol{T}_{i;B} \tag{6.32}$$

However, as Kim takes the definition to require the property sets be closed, we should write this last relation as:

$$\overline{\boldsymbol{T}}_{i;A} \looparrowleft \overline{\boldsymbol{T}}_{i;B} \tag{6.33}$$

Kim (1993) goes on to propose a second definition of weak supervenience:

> A *weakly supervenes* on B if and only if necessarily for any property F in A, if an object x has F, then there exists a property G in B such that x has G and if any y has G it has F. (Kim, 1993, p. 64)

[24]Many of these are collected in Kim (1993).

Now this is only equivalent to the first definition if we take it that the property set are Boolean complete, as otherwise there is no guarantee that property G exists at all. Recall that the Boolean completion allows every set of indistinguishable objects to be associated with a unique property. However, there is still a subtle difference between our definition and Kim's. He does not mention the object sets that are involved in the supervenience relations and implies that the object set is actually all possible objects in the world in question. For our definitions this is not a trivial matter as it may be the case that two property sets do supervene in one manner under our definition but fail to do in that manner so under Kim's. This would be the case if the objects are restricted in our pre-topologies [or topologies] in comparison to possible sets of objects to which the properties could be applied. We shall take it that Kim's definitions involve all the objects in a world at a specific time.

Kim (1993) goes on to define strong supervenience as one that has some trans-world stability and force:

> A is said to *strongly supervenes* on B just in case:
>
> necessarily, for any object x and each property F in A, if x has F, then there is a property G in B such that x has G, and *necessarily* if any y has G, it has F. (Kim, 1993, p. 65, p. 80)[25]

Before looking at this in our terminology a number of comments are useful. First, Kim intends the property sets to be Boolean complete, which he states explicitly, and the object set to be complete in all worlds under consideration. Secondly, the modal force of the two uses of 'necessarily' and Kim allows the precise usages to remain undetermined in the definition: Kim (1993) says,

> It is neither possible nor desirable to specify in advance how necessarily is to be understood here; an appropriate specification must depend on the particular supervenience thesis under consideration and different readings of "necessarily" will yield different supervenience theses to consider. (Kim, 1993, p. 66)

Kim goes on to endorse McLaughlin's definition of supervenience to be equivalent to his definition above. Kim (1993) quotes McLaughlin's definition of strong supervenience:[26]

> For any worlds w_j and w_k, and for any objects x and y, if x has in w_j the same B-properties that y has in w_k, then x has in w_j the same A-properties that y has in w_k. (Kim, 1993, p. 81)

[25]The quote here is from page 65. The definition on page 80 has slightly different wording.

[26]Kim notes that the quote comes from an unpublished manuscript. The manuscript is entitled, "Why Try to Bake an Intentional Cake With Physical Yeast and Flour?".

Which we can write as follows

$$\overline{\boldsymbol{T}}_{i;A} \looparrowright_S \overline{\boldsymbol{T}}_{i;B} \Leftrightarrow x(w_j) \sim_{i;B} y(w_k) \Rightarrow x(w_j) \sim_{i;A} y(w_k) \quad \forall w_j, w_k \tag{6.34}$$

Again, as previously, this definition requires both Boolean completion of property sets and maximal object sets and some notion of sameness of objects across worlds. We could, as for weak supervenience, relax the Boolean completion, but we shall stick with Kim's definition as otherwise the different definitions come apart from each other.

6.2.7 Supervenience across different object sets

We can extend our idea of supervenience further by allowing property sets to operate on different object sets. The motivation behind this is threefold: it will allow us to construct general definitions that deal with how supervenience relates to the way properties are distributed across objects in different worlds - something we examine later in this section; and also it will furnish us with the technical apparatus to deal with supervenience of object sets as they evolve or alter in time, should we wish to do so. Additionally, it is a natural mathematical step to look at generalising the simple versions of supervenience we have outlined in earlier sections.

In this section, then, we set out how to combine supervenience with maps between different object sets and briefly illustrate how this can be used to define some more sophisticated versions of supervenience that appear in the literature. We shall also need to invoke the techniques we outline here in some aspects of the discussion we present below. We leave a detailed development of these techniques, however, for another time and place.

If we are to deal with more than one set of objects - one for each property set - then, in order to maintain the spirit of the definitions we have set up, we will need to have some way of associating the objects in one pre-topology with the objects in another. To this end we shall define a map between object sets and then use this map to produce some more general definition of supervenience. For convenience we shall take each object set as being formed of objects within one world only, although these worlds may differ between the object sets themselves.

Given two object sets,[27] $\mathbb{O}_i = \{O_p(t, w_i); p \in \mathcal{I}\}$ and $\mathbb{O}_m = \{O_q(t, w_m); q \in \mathcal{I}\}$ then we take a

[27]We suppress the spatial dependence, Δx, that we had in our original specification; and we ignore issues with trans-world notions of time.

bijective map f from the objects of the first set to the objects of the second set:

$$f : \mathbb{O}_i \to \mathbb{O}_m \tag{6.35}$$

$$f(O_p(t, w_i)) = O_q(t, w_m) \tag{6.36}$$

We require that the map f is always a bijective map, recalling this means that every object in $\mathbb{O}_m = \{O_p(t, w_m); p \in \mathcal{I}\}$ is associated with a unique object $\mathbb{O}_i = \{O_q(t, w_i); q \in \mathcal{I}\}$ and vice versa. Requiring the map between object sets to be bijective means we can only define our general supervenience when we have object sets with the same cardinality. We note that our map relates objects at the same time but, possibly, at differing worlds. We could change this and allow the map to relate objects in the same world at differing times.

In order to distinguish supervenience between pre-topologies [or topologies] with different object sets from those between pre-topologies [or topologies] with identical object sets, we shall call the former **general supervenience with respect to f** and the later **supervenience** in keeping with our use so far. The other definitions we have set up will carry across to general-supervenience naturally.

We define two pre-topologies as generally-supervening with respect to f - referring to (6.36) - as follows:

$$\boldsymbol{T}_{i;k} \underset{f}{\hookrightarrow} \boldsymbol{T}_{m;n} \Leftrightarrow (f(O_j) \sim_{m;n} f(O_l) \Rightarrow O_j \sim_{i;k} O_l) \tag{6.37}$$

And noting that we could define supervenience as a special form of general supervenience by making use of the Identity map, Id, between an object set and itself, mapping each object to itself:

$$\boldsymbol{T}_{i;k} \underset{Id}{\hookrightarrow} \boldsymbol{T}_{i;n} \tag{6.38}$$

Instead we choose to suppress the appearance of Id and stick with our original notation.

We can use this idea of general supervenience to formally construct a way of stating when two pre-topologies, viewed as groupings of objects induced by their respective property sets, are in some sense identical or mirrors of each other. Essentially we require that the properties divide up the objects into sets in the same way in each pre-topology so that if two objects are indistinguishable in one pre-topology then their images are indistinguishable in the other pre-topology. We can write this condition using the distance function defined earlier. Recall for objects in a pre-topology $\boldsymbol{T}_{i;k} = \{\mathbb{O}_i, \Pi_k\}$ we defined the distance between two objects using

$$\mathfrak{d}_{i;k}(O_j, O_l) \equiv \; max\|P_m(O_j) - P_m(O_l)\| \;\; \forall P_m \in \Pi_k \tag{6.39}$$

Using this, we can say that two pre-topologies are f-isomorphic, which we denote as $\boldsymbol{T}_{i;k} \equiv_f \boldsymbol{T}_{m;n}$ with respect to a bijective map $f : \mathbb{O}_{i;k} \rightarrow \mathbb{O}_{m;n}$, if:

$$\mathfrak{d}_{i;k}(O_j, O_l) = \mathfrak{d}_{m;n}(f(O_j), f(O_l)) \;\; \forall O_j, O_l \tag{6.40}$$

Notice that two given pre-topologies may be f-isomorphic under one bijection but not under another. If there is some f for a given pair of topologies then this tells us that the property sets divide each of their respective object sets up in the same general way. As such the existence of an f-isomorphism between pre-topologies [or topologies] furnishes us some general information about the property sets. We can use this idea of f-isomorphism to define a more general idea of isomorphism of pre-topologies by requiring there to be some f such that they are f-isomorphic. As f-isomorphism is symmetric, reflexive and transitive this sets up an equivalence relation amongst pre-topologies.

If we allow both pre-topologies to have the same set of properties the definition of f-isomorphism does not need to preserve the property behaviour between the topologies, it only requires that the way objects are grouped is unchanged by the mapping even though the grouping may be achieved in different ways in each pre-topology. If we want to preserve the way the properties achieve the groupings of objects under the mapping as well then we need to restrict our definition slightly. We shall define a property preserving isomorphism [which we shall label a p-isomorphism] between pre-topologies as one that is both an f-isomorphism and which also preserves property values. So for each property P_m in the relevant set, we additionally require:

$$P_m(O_j) = P_m(f(O_j)) \tag{6.41}$$

As mentioned, the p-isomorphism is somewhat more restrictive than an f-isomorphism. The advantage, perhaps, of an f-isomorphism over a p-isomorphism is that it suggests there may be some deeper potential relationship between diverse sets of properties and allows some form of analogies to be drawn.

We can use the definition of p-isomorphism to reproduce the definitions of weak and strong global supervenience as they appear in the literature. To ensure our definitions reflect those in the literature we must take it that each of our object sets in our pre-topologies are world-bound, that is they each exist in one world only - although these worlds may differ between object sets - and are exhaustive in that they contain all the objects of that world. We must also take it that we can only talk of worlds where this is the case. For convenience, we shall take it that the index of the object set also labels the corresponding world, so set $\mathbb{O}_a$ is in world w_a, and as we are dealing with p-isomorphisms we must remember that the property set for each pre-topology is the same. As an example, we can set out two definitions as they appear in the literature, using our notation. First, we look at a characterization of two types of supervenience as they appear in McLaughlin (1997):

> "... we should follow the common practice in mathematics of appealing to the notion of there being some *isomorphism*, without actually specifying any isomorphism. Let us say that an isomorphism, I, between two worlds w and w* *preserves* B-properties if and only if, for any x in w, x has a B-property F in w if and only if the image of x under I has F in w*. We may say, then, that two worlds have the same *world-wide* pattern of distribution of B-properties just in case there is some isomorphism between them that preserves B-properties. It follows that two worlds can have the same such pattern only if they have the same cardinality. But that is what one would expect from the notion of two worlds having the same world-wide pattern of distribution of B-properties. Employing this notion of a world-wide pattern of property distribution, notions of strong and weak global property supervenience can be defined as follows:
>
> *Strong Global Property Supervenience.*
> A-properties globally supervene on B-properties = df for any worlds w and w *, every B-preserving isomorphism between them is an A-preserving isomorphism.
>
> *Weak Global Property Supervenience.*
> A-properties globally supervene on B-properties = df for any worlds w and w *, if there is a B-preserving isomorphism between them, then there is an A-preserving isomorphism between them.
>
> Suffice it to note that it is sometimes uncertain which of these two notions is intended when one finds global supervenience invoked in the literature."
> (McLaughlin, 1997, p. 214)

We can translate these as follows:

DEFINITION 6.4: Property set Π_A **strongly globally supervenes** [at time t][28] on property set Π_B if and only if for any two worlds, w_j and w_k, if $\boldsymbol{T}_{j:B}$ is p-isomorphic to $\boldsymbol{T}_{k;B}$ with respect to a bijection f then $\boldsymbol{T}_{j;A}$ is p-isomorphic to $\boldsymbol{T}_{k;A}$ with respect to the same bijection.

DEFINITION 6.5: Property set Π_A **weakly globally supervenes** [at time t] on property set Π_B if and only if for any two worlds, w_j and w_k, if $\boldsymbol{T}_{j;B}$ is p-isomorphic to $\boldsymbol{T}_{k;B}$ then $\boldsymbol{T}_{j;A}$ is p-isomorphic to $\boldsymbol{T}_{k;A}$

As we mentioned above, we leave the application of our techniques at this point. We hold that what we have introduced is sufficient to set out much, if not all, of the definitions and proofs that appear in the literature in a formal manner. We leave the details of this for another place.
In what follows, we shall concern ourselves with supervenience only within a single world and so have little more to say on trans-world supervenience. The discussion in Chapter 4 and the fact that the model we constructed for free will in Chapter 5 is world-bound provide the rationale behind this restriction to a single universe henceforth.

6.3 Supervenience and causality

6.3.1 Introduction

Recall again some of the aims we set ourselves at the outset of this study: to develop mathematical models and to use them to examine free will. In this section we aim to combine both the work on causality we set out in Chapter 3 with the models we constructed above to deal with supervenience. We do this partly in fulfilment of our aim to produce mathematical models and partly in our pursuance of free will.
In what follows, then, we set out how we might understand properties and objects in terms of our causal relata. We start by providing some motivation for our unpacking of objects and properties in terms of relata. We set up the details in such a way that we can segue into talking of supervenience and causality in the same breath. This will allow us to examine how they fit together.
In fulfilment of one of our other aims - the pursuance of an ontic version of free will - we shall then examine our earlier arguments in the light of this melding between supervenience and causality. We examine our arguments to determine whether considerations of supervenience might suggest

[28]We insert this phrase here merely to emphasize our discussion in section 6.2.4.

that our approach is in some manner incomplete or incorrect. We conclude that our arguments stand.

Before we begin: in what follows we confine ourselves to basic single-world supervenience. We do this because it is adequate for our needs but also because we have outlined, in Chapter 4, some suspicions as to the use of modal concepts when dealing with ontic restrictions to universes.
And further to our comment in section 6.2.4, we explain how we are to unpack supervenience in this section. First, we assume that general properties are time independent but that our property sets are not necessarily Boolean complete - we say more on properties sets in sections 6.3.3 and 6.3.5 below. Second, we take it that our object sets are the complete sets of objects extant within a universe at a specific time. Third, we extend our time specific definition of supervenience to be time free in this section. We achieve this as follows: given two property sets Π_a and Π_b, we say that former supervenes on the later if and only if:

$$\{\mathbb{O}_i(t), \Pi_a\} \looparrowright \{\mathbb{O}_i(t), \Pi_b\} \ \ \forall t \tag{6.42}$$

Where the object sets are taken to be the complete set of objects extant at the specified time. In other worlds we take it that supervenience of one property set on another must be examined with respect to every complete object set at every individual time within a universe. We note that, should we wish to modify the times for the object sets in (6.42), we could invoke the techniques we set out in section 6.2.7 above. What (6.42) allows us to do is to put aside worries as to object persistence and the ambiguities that we discussed in section 6.2.4 above.

As supervenience is a relationship between property sets with respect to some specified object set, we shall start by unpacking both what are to be taken as objects and what are to be taken as properties within our causal structures.

6.3.2 Objects

What are to be our object sets? When discussing of the supervenience of the mental on the physical in section 6.2, we took our objects to be, loosely, brains or some such. And our conception of objects tends to involve things such as tables, chairs, dogs, cakes and the like. In all these cases our objects are very specific temporally persisting lumps of matter occupying well-defined, or fairly well-defined, spatial regions: in other words things like, brains, apple pie, cats and the like. More generally it seems intuitively appealing to define an object as the material occupant of some well-defined spatial region and talk about the same object at different times by following this lump of matter as it wends its way across space. This sort of definition of an object is, however, difficult to reconcile with a number of factors.[29] For instance, there are many problems with this approach to objects - as material occupants of spatial regions - that stem from physical theories. There is a problem with fundamental particles - such as quarks, leptons etc - considered within the standard model of particle physics:[30] they are considered point particles in the theory.[31] Equally, it is not clear what is to be done with entangled states, such as entangled pairs of electrons, in that they might be considered to constitute a single object even though they have spatially disjoint 'parts'. And then what about bosons that may not have individual identity in any clear sense, as discussed by Teller (1997)? And, what then of spontaneously created particle-antiparticle pairs? What is more, even in our everyday talk we do not usually consider objects to be the full content of any spatial region. For instance, if I wish to talk about one billiard ball's hitting another and causing it to fall into a pocket, then I tend to eschew a specification of the billiard ball as being the full contents of a closed, simply-connected spatio-temporal region. For instance, I tend to ignore any fluctuating electromagnetic and gravitational fields inside its boundary. These internal fields may be altered by earlier movements of planets, the broadcast from the local radio station, remnants of the big bang and so on. Should we include these in the full description of a billiard ball and so include the big bang as part of the detailed cause of the billiard ball's falling into a pocket? The idea seems absurd.[32] Rather, at least in causal cases, we isolate in some *ad hoc* manner what is relevant and ignore what is deemed irrelevant. This suggests that we should not take the complete content of closed, simply-connected spatial regions as being in some sense representative of what we mean by objects, assuming our conception of object has some ontic purchase. Rather we propose to define our objects as being constituted of relata, as sets of fundamental relata, and then define any notion of persistence of objects through time using causal concepts. We take it that not all conglomerations of fundamental relata constitute objects but, rather, that there is some ontic restriction on which relata are objects.

In defining objects this way we face two problems: the first is that whilst we want to eschew the idea of spatial locality as a defining aspect of objects in some cases, we should hope that any conception of object we propose would furnish us, mostly at least, with the sorts of things we do

[29]For instance, see discussions in Poland (1994) and Markosian (2000).

[30]The standard model is the mathematical and interpretive physical theory that deals with elementary particles. For introductory discussions see, Feynman & Weinberg (1999) and Griffiths (2008).

[31]String theory has a different take on these and gives them spatial extension in a certain manner; see Zwiebach (2004).

[32]Despite seeming absurd, it is important to note that such causal relations may well be the case according to the schema set out in Chapter 3. For instance, it may be the case that electromagnetic fields cannot be considered as separate relata in some cases and so must form part of, for instance, a billiard ball. And then the full cause of a billiard ball's falling into a pocket would include such things as the broadcast from the local radio station. Of course the hope would be that our intuitive understanding of the way the world works and what is relevant and irrelevant to certain causal processes are faithful mimics of the ontic causal structure.

recognise fairly well as objects - billiard balls, pieces of cake, trees, planets and so forth. That is, we would hope that the collections of relata that are our objects will have some spatial element to them. Why is this a difficulty for our definition of object? When causal relata were introduced in 3.2 we took them to be a complete description of all that there is. Now if we take it that the spatial manifold is in some sense fixed as a background in which these relata are placed then they can inherit their relative localities from the localities of the manifold. However, this then brings into question the causal role of the manifold itself, as it must have ontic presence and temporal persistence too and a causal role to play. In response to this difficulty we proposed that the juxtaposition of relata, that is the very idea of space, is more likely to arise from the way relata themselves interact from moment to moment. That is to say that any manifold we might wish to use to talk of relata juxtaposition must itself be a facet of the causal process in some manner. This provides us with a difficulty as we have not said in what manner the manifold is to arise from the relata: certainly no easy task.[33] So we must admit that we have no basis on which to move from our definition of an object as a set of relata to an understanding of an object as generally spatially well-defined and local. This is a lacuna in the model but one we suspect we must live with for the present and additionally one that does not affect the arguments we advance when considering supervenience below.

The second issue which may provide us with some difficulty was discussed in 2.4. In essence, we asked there how we are to reconcile discrete time and our discrete universe states with the idea of a persisting object. Again, we leave this matter unresolved here and allow ourselves to talk of objects at different times without worry.

These issues are important and difficult to resolve easily.[34] However, if we merely define our objects as certain relata, that is certain [structured] sets of fundamental relata, at any given moment then our definition will be adequate for our needs here. We shall assume there is some way of unpacking and defining an idea of persistence through time and locality but, when we need such ideas, we shall revert to the intuitive concepts we have of everyday objects, such as brains and billiard balls, and ignore difficulties of entangled states, point quarks and so on.

To summarise: we take our objects to be relata, that is collections of fundamental relata. However, it is not the case that any collection of fundamental relata extant at some point in time will

[33]In fact it is a very difficult task and we do not know of any completely successful attempts. There is some indication that space and time arise from something more fundamental in such areas as loop quantum gravity, see Rovelli (1998), as well as suggestions that this might be an interesting approach to take in the use of non-commutative differential geometry in trying to grasp quantum theory - see, for instance Connes (1994).

[34]At present we do not know how to resolve them satisfactorily.

constitute an object. Rather, we shall assume there is some ontic restriction on what groupings of fundamental relata constitute objects and that this restriction does reflect in some manner, albeit vague, the locality and everyday conception of objects we have as being occupiers of space in some manner. We shall also assume that more exotic candidates for objects, such as entangled states, bosons and the like are catered for under our definition and leave the details for another time. And to reiterate our discussion above: we take our sets of objects as comprising *all objects* in the universe at each given moment - one set for each moment.

6.3.3 Properties

Here we shall define properties simply as causal aspects of relata, or of fundamental relata, and leave it at that. This is, of course, a little vague but what we can say, however, in mitigation towards plausibility, is that we usually infer properties of things from the way they interact causally with us directly or with other objects in the universe. So, for instance, the deduction there are electrons with charges and spin and the like derives from a combination of the observation of regular patterns in the world and mathematical models,[35] and of course a lot of additional conceptualization as well. Here we take our idea of property from the fact that properties concern how things affect each other and we take it that this gives some, albeit rough, justification for our definition.

In order to get some purchase on supervenience across time we shall make the assumption that properties have trans-temporal ontic validity and a notion of sameness obtains: so we could talk about a property type without reference to time should we wish. However, it is important to note that in what follows many of our definitions will start with specific instances of properties and objects and then move to general definitions. That is we countenance the idea of instance of properties [tokens] and the idea of property in general [types] without further comment or debate. So we can talk, for instance, of a specific electron having a charge of +1 or the general property of having a charge of +1.

As a final point, it is worth noting that the relationship between fundamental relata and properties is not simple. We have said above that fundamental properties are essentially the causal parts of fundamental relata. However, the causal parts of fundamental relata are set by both the content of the relatum *and* the ontic laws of nature that the relatum is subject to. If it is the case that all

[35] Noting that this rest on the implicit assumption that the world represents itself to us in some [mathematically based] consistent manner from moment to moment and that this representation is in turn is some sort of consistent reflection of the way the world is, at least in part.

relata are essential, and recall that essential relata held their causal nature within their content, then it is true to say that content-identical relata will have identical properties too. However, if we countenance separable relata, recalling that this is altogether different from countenancing separable universes, then it may well be the case that fundamental relata can be content-identical and yet have differing properties or, alternatively, differing causal effects.

6.3.4 Causal equivalence for properties

In this section and the next we start our move towards explaining how supervenience sits with causality. Specifically, here we start to look more closely at how properties, objects and causality fit together. This will allow us to unpack various ways of interpreting supervenience - particularly how it might sit with reductive relationships between sets of properties.

First, it is useful to recall that supervenience is a relationship between sets of properties and not a relationship between the objects themselves. In fact, as we emphasized earlier, the object sets serve only to identify the way the properties in question group together. Above we suggested that we should not consider objects as being exhausted by the content of closed, perhaps simply connected, spatial regions but rather we should think of them as sets of relata. Here it is worth adding the note that we do not have to worry too much if objects are merely conglomeration of properties or are perhaps something else on which properties can be hung in some manner. Rather we need only consider objects in so far as they represent actual, that is ontic, conglomerations of properties.

As we have taken properties to be causal aspects of relata we have a further set of relationships between properties, namely causal relations via their associated relata. We now explain how we shall deal with this formally.

We hold, simply, that if two relata are causally related then their properties can be said to be causally related too. Recalling that fundamental relata are causally complete then we can define a simple map, Φ that takes a specific fundamental relatum to its corresponding property instance:

$$\Phi : f\mathcal{S}_i(t_a) \to p^i(t_a) \tag{6.43}$$

Noting that we indicate property instances using superscripts and general properties - appearing in our discussions of supervenience above - using subscripts. This convention is used throughout this section. When the context is clear, we use the term ‘property’ for both.

We assume each fundamental relatum has only one property associated with it and we shall call these fundamental properties - denoted by lower case ‘p’. We combine this map with the causal maps we defined in Chapter 3 so that we can talk naturally of causal relations between property instances in the same manner we talked about causal maps between fundamental relata earlier. So we can write:[36]

$$\Phi(\mathfrak{C}(t_a, t_{a-1})(f\mathcal{S}_i(t_a))) = \Phi(f\mathcal{S}_j(t_{a-1})) \tag{6.44}$$

and we abbreviate this to:

$$\mathfrak{C}(t_a, t_{a-1})(p^i(t_a)) = p^j(t_{a-1}) \tag{6.45}$$

or even:

$$\mathfrak{C}(p^i(t_a)) = p^j(t_{a-1}) \tag{6.46}$$

This is, in fact, a slight simplification as often the cause of a fundamental relata is a set of earlier relata. In this case we can write:

$$\mathfrak{C}(p^i(t_a)) = \{p^j(t_{a-1})\} \tag{6.47}$$

Where the brackets indicate a set of properties is involved. Of course, the right-hand side of (6.47) does not indicate the property set is associated in any way with a single object.

The map naturally extends to causal relationships between properties and it allows us to flit between property instances and relata when talking of causality with little further comment. This allows us to bring the causal machinery we developed earlier to properties. Henceforth, we shall assume the Φ map is in place in what follows without explicitly mentioning its presence.

We can now state the condition for two *instances* of contemporaneous fundamental properties to be **causally-equivalent**, denoted by $\doteqdot$, at some time t. In simple terms, two properties [and hence, fundamental relata] are causally equivalent if they both have exactly the same specific effect on the future:

[36] For ease of exposition we have assumed the output of causal maps here is a single fundamental relatum rather than the more usual [structured] set of fundamental relata.

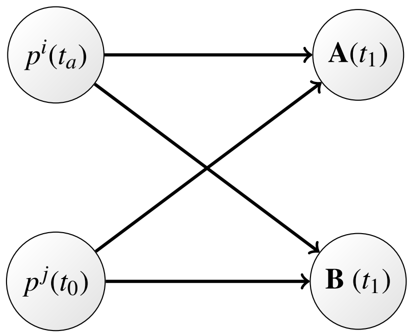

Figure 6.1: Symmetric causal over-determination as causal-equivalence

DEFINITION 6.6a: Causal equivalence of properties

Two properties, $p^i(t_0)$ and $p^j(t_0)$, are **causally equivalently**, denoted $p^i(t_0) \doteq p^j(t_0)$, if both,

$$\mathfrak{F}(t_0, t_1)(p^i(t_0)) = \mathfrak{F}(t_0, t_1)(p^j(t_0)) \tag{6.48}$$

and for each $p^k(t_1) \in \mathfrak{F}(t_0, t_1)(p^i(t_0))$

$$\mu(p^i(t_0), p^k(t_1)) = \mu(p^j(t_0), p^k(t_1)) = \mu(\{p^j(t_0), p^i(t_0)\}, p^k(t_1)) \tag{6.49}$$

Where we use $\mathfrak{F}(t_0, t_1)$ from section 3.7.

We can also think of this definition as saying that two properties both existing at the same time are **causally-equivalent** if and only if they have the same causal consequence as token instances - as in (6.48) - and they both contribute in exactly the same way to their causal consequences - as in (6.49).

It might seem that the properties instances that are causally equivalent must be the very same property instance, and hence be underpinned by a single relatum, but this is not the case. It is possible for two properties to be symmetrically over-determining, and causally equivalent without their being the very same property instance. This is illustrated in Figure 6.1. In Figure 6.1, we could remove either $p^i(t_0)$ or $p^j(t_0)$ without affecting the content of either $\mathbf{A}(t_1)$ or $\mathbf{B}(t_1)$.

However, if we consider only situations where there can be no causal over-determination, then any properties that are causally-equivalent are really just the very same single property. This is an important result as it shows that properties can be causally indistinguishable and yet still remain distinct so long as causal over-determination is allowed.

We can extend causal equivalence to contemporaneous relata and to their corresponding properties by building up instances of properties from the union of the properties derived from the relatum's constitution by fundamental relata. Before we do so we need to make some comments on how properties are to be combined here.

Recall, we have defined objects as certain relata, that is certain structured sets of fundamental relata. Not every structured set of fundamental relata is to be an object and we have assumed that there are some, albeit unspecified, restrictions, on which structured sets are to count as objects. As such, if we take some object, $O_i(t_a)$ as a structured set of fundamental relata:

$$O_i(t_a) = \{f\mathcal{S}_j,\ some\ \ j\} \tag{6.50}$$

Then we may associate some property, P^i with $O_i(t_a)$ and define it in terms of the properties of some of its constituting fundamental relata as follows:

$$P^i(t_a) = p^k \cup p^l \ldots \tag{6.51}$$

where the properties p^k etc on the right-hand side are the properties derived from the fundamental relata appearing in $\{f\mathcal{S}_j,\ some\ \ j\}$ under the map Φ defined in equation (6.43) above. Further, the right-hand side of (6.51) need not be composed of all the properties associated with the object in question, a subset is sufficient to form *some* property associated with $O_i(t_a)$.

In what follows, we shall make one simplifying assumption concerning property unions and intersections. We shall assume that the right-hand side of (6.51) is unaffected by the *structured* nature of the set of fundamental relata that constitute the corresponding object. Were we not to do this then we would have to take account of the possibility that objects made from the same *set* of fundamental relata but in differing structural arrangements could have differing properties - the P^i arising from the same set of base properties - the p^j - and hence we would need to take account of this when defining our terms. This would add an extra structural complication to property combinations and, whilst we suspect this may well be necessary in more complete treatments, it would considerably complicate the discussions that follow and the definitions of property combination. We note that no such structural restrictions appear in the literature on supervenience when Boolean combinations of properties are considered - we consider this as a potential weakness both of the literature and of our treatment here. Nevertheless, we hold without comment that this weakness does not detract from our aim here - to argue that considerations of supervenience do not affect our causal arguments in any significant way.

We then have a restriction on the combination of property instances: we allow countable set theoretic operations for property instances so long as the property sets under consideration are associated with a single object - or object type - and we do not carry across the structured aspects of relata to the sets we use when considering properties.

We now return to our theme: we can extend the definitions of causal equivalence above to these relata. So, for instance if $P^i = p^a \cup p^b$ and $P^j = p^c \cup p^d$ then we can write:

$$(p^a \simeq p^c) \wedge (p^b \simeq p^d) \Rightarrow P^i \simeq P^j \tag{6.52}$$

or

$$(p^a \simeq p^d) \wedge (p^b \simeq p^c) \Rightarrow P^i \simeq P^j \tag{6.53}$$

In the reverse situation, defining causal equivalence of properties in terms of their fundamental constituents, we have to be a little more careful as it is possible that two properties may have some fundamental properties in common and some that are distinct but causally-equivalent.
If we are given two properties, $P^i(t_a)$ and $Q^j(t_a)$ then we first decompose them into a set of fundamental properties:

$$P^i(t_a) = \{p^\alpha, p^\beta ... p^\theta\} \quad Q^j(t_a) = \{q^\alpha, q^\beta ... q^\phi\} \tag{6.54}$$

Noting that there may be a different number of fundamental properties in each set - as some may be causally-equivalent to others, we have the following definition:

DEFINITION 6.6b: Causal equivalence of properties

$$P^i(t_a) \simeq Q^j(t_a) \Leftrightarrow (\forall p \in P^i, \exists q \in Q^j : p \simeq q) \wedge (\forall q \in Q^j, \exists p \in P^j : q \simeq p) \tag{6.55}$$

And we shall add that they are **distinct** and causally-equivalent when they have no fundamental property instances in common:

$$P^i \cap Q^j = \emptyset \tag{6.56}$$

Where, as we mentioned, all set operations must be justified by being related to objects.
We note in passing that if we disallow the possibility of over-determination, then any two properties that are causally-equivalent must be the very same property.
So far we have dealt only with specific token instances of properties. In the next section we extend this to properties in general - types - and explain how this allows us to categorise a number of ways of understanding how supervenience may arise.

6.3.5 General property operations

In this section, we make some points about the differences between property instances and the properties we use in our definitions of supervenience. This will clarify the scope and assumptions we shall make in following sections.

Above we allowed for certain set theoretic operations on property instances associate with specific objects. We held that all set operations for these object bound properties must take place with the underlying object as defining the scope of properties that could be considered for some set theoretic operation. And further, recall, we made the decision to eschew the possibility of structure coming into play when property instances were combined set-theoretically.

Now we extend set-theoretic combinations to general properties. However, general properties are somewhat different in their mathematical and ontic status from the object bound property instances we discussed above. General properties are here to be considered as maps that map objects to either 0 or 1 depending on whether the object in question possess an instance of the general property under consideration. We defined the operations of combinations of these maps in equations (6.16), (6.17) and (6.18)

For example, consider the object O_i constituted of a set of property instances, $\{p^a, p^b, p^c\}$. That is we set:

$$O_i = \{p^a, p^b, p^c\} \tag{6.57}$$

Now if we wish to state whether this particular object possesses an instance of p_a we can use a general property considered as a map. We denote general properties using subscripts and property instances using superscripts.[37] Thus we denote the map that tells us whether an object possesses p^a by "p_a". So we write:

$$p_a(O_i) = p_a(\{p^a, p^b, p^c\}) = 1 \tag{6.58}$$

The result is 1 because the object does possess an instance of p_a .

It is also worth noting that property instances are time and position bound as each derives from a specific fundamental relatum according to (6.43) above. In contrast, general properties are not time or position bound - and we have also assumed, in setting out our model of supervenience above, that they have a trans-world presence.

We can now clarify an important and subtle point concerning the set-theoretic combination of gen-

[37]We did *not* use any such convention with fundamental relata as each is considered unique: its relative 'position' and time of occurrence are also relevant.

eral properties. First we note that supervenience concerns sets of general properties and sets of objects that possess property instances. We might take our object set as being all objects present at some specific time or all objects that exist, regardless of the time span of their existence. However, what is important here is the subtle difference between the way we can effect set-theoretic operations for individual property instances and how we effect set-theoretic operations for properties in general. Recall, that we only permitted specific properties to be combined set-theoretically within [the scope of] an object. We place no such restrictions on general properties and we permit all possible countable set-theoretic operations on general properties. We do this primarily to keep our structures in line with those in the literature where Boolean completion plays an important role - although, as we mentioned above, we do not assume our property sets are Boolean complete. What we must ask, however, is what is the significance of this decision. The answer is straightforward: there may well be combinations of general properties which are not ever reflected within a single object - and, note, this cannot happen to set-theoretic compilations of property instances. For example, consider the toy universe constructed from two objects:

$$O_1 = \{p^a, p^b, p^c\} \qquad\qquad O_2 = \{p^d, p^e\} \tag{6.59}$$

And consider the set of general properties:

$$\Pi = \{p_a, p_a \cup p_b, p_a \cup p_d\} \tag{6.60}$$

Certainly, general properties p_a and $p_a \cup p_b$ appear [as property instances] in some objects, but $p_a \cup p_d$ has no presence within a single object. These then gives some interesting results when considered as mappings:

$$(p_a \cup p_d)(O_2) = 1 \tag{6.61}$$

even though O_2 does not possess both p_a and p_d. Whereas,

$$(p_a \cup p_b)(O_1) = 1 \tag{6.62}$$

We see that in (6.61) O_2 does not possess both properties whereas in (6.62) O_1 does possess both p_a and p_b.

That is, the mappings and the idea of property possession come apart a little if any set-theoretic combination of general properties are included. They come apart because results such as (6.61) and (6.62) do not differentiate between subtle differences in property possession and they also come apart because, in some cases, no object possess the complete property achieved by set-theoretic operations - as in the case of $p_a \cup p_d$.

We shall not worry about this too much but it shows, along with the lack of structural consideration, a weakness in blithely allowing Boolean completion for general properties in supervenience definitions.

What is important for what follows is that we shall allow general property combination without concerning ourselves with whether these sorts of problems arise. A more complete study would take account both of the schism between set theoretic operations for general properties and property instances and also take account of the possible need to incorporate structural aspects into property combinations. We shall return to this briefly in section 7.3.3 where we make some points concerning supervenience and the mind-body problem.

6.3.6 Modes of supervenience

In this section we shall outline a number of ways that the supervenience relation can arise between sets of properties.

We shall start with a definition of what it is for one property instance to reduce to others.

> **DEFINITION 6.7: Property instance reduction**
>
> A property instance $P^i(t)$ **reduces** to a set of properties $\{q^i(t), q^j(t), ...q^n(t)\}$ if and only if $P^i = q^i \cup q^j \cup ... \cup q^n$ at t.

and we could extended this to properties in general:

> A property $P_i(t)$ **reduces** to a set of properties $\{q_i, q_j, ...q_n\}$ if and only if for every instance of the property we have $P^i = q^i \cup q^j \cup ... \cup q^n$.

However, we note we have couched our definition in such a way that reduction is parasitic on the way properties appear in individual objects. In light of the discussion in section 6.3.5 above, this could prove to be problematic if there is a schism between the set-theoretic operations for property instances and those for general properties. So we define a theoretical reduction of general properties without reference to individual objects or times, as follows:

> **DEFINITION 6.8: General property reduction**
>
> A [general] property P_i **reduces** to a set of properties $\{q_i, q_j, ...q_n\}$ if and only if $P_i = q_i \cup q_j \cup ... \cup q_n$

It is this, latter, definition we shall use.

Further, we extend the definition to one set of general properties reducing to another by requiring each property in the reduced set to be expressible as the union of some subset of properties in the reducing set and, further, that each member of the reducing set appears in at least one reduction of a member of the reduced set.[38]

Now, before moving to look at how supervenience is connected with causality, we need to extend the definitions of causal equivalence to general properties considered at some time. Recalling that we denote proprieties in general at some given time using subscripts. We construct our definition as follows:

> **DEFINITION 6.9: Time specific causal equivalence**
>
> A P_i is **causally equivalent to** Q_j **at time** t_a if for every instance of P_i there is an instance of Q_j such that $P^i \simeq Q^j$.

> **DEFINITION 6.10: Causal equivalence**
>
> A P_i is **causally equivalent to** Q_j if P_i is causally equivalent to Q_j at all times when P_i occurs

In contrast to reducibility above, we have left the definition of causal equivalence for general properties parasitic on property instances. We also note that 'causally equivalent to' is not necessarily a symmetric relationship. Recall that a symmetric relationship is any binary relation, R, that is indifferent to the order in which the variables appear. So if R is symmetric we can write $\forall a, b : aRb \Leftrightarrow bRa$. Here that is clearly **not** the case for causal equivalence. So we shall write $P_i \triangleright Q_j$ for our second definition - and we note $P_i \triangleright Q_j \not\Leftrightarrow Q_i \triangleright P_j$.

We then extend the idea to sets of properties:

> **DEFINITION 6.11: Causal equivalence for property sets**
>
> $\Pi_1 = \{P_i\}$ is **causally equivalent to** $\Pi_2 = \{Q_i\}$ if and only if $\bigcup_i P_i \triangleright \bigcup_i Q_i$

Our definitions of **causally equivalent to** may not seem the most natural but they are constructed specifically to be useful in the discussion of pure-supervenience and causality that follows below. Now consider two sets of properties Π_i and Π_j such that $\{\mathbb{O}, \Pi_i\} \hookleftarrow \{\mathbb{O}, \Pi_j\}$, where our object set is the full set of objects [relata] extant at each time and supervenience is defined over these as per (6.42). And note that our property sets are made up of properties in general.

[38] And noting that if one set of properties instances reduces to another then the two sets must be causally equivalent; but causal equivalence, of course, does not mean reduction if over-determination is countenanced.

The following are basic possible ways in which supervenience can arise when $\{\mathbb{O}, \Pi_i\} \looparrowright \{\mathbb{O}, \Pi_j\}$:

[Sup1] Simple Reductive Supervenience

Each member of Π_i reduces to a set of fundamental proprieties which is itself the union of one or more of the sets, or a subset of one or more of the sets, that constitute the reduction of properties in Π_j to fundamental property sets.

Before moving to the two other modes of supervenience, we provide a simple example of **[Sup1]**:

$$O_1 = \{p^a, p^b\}$$
$$O_2 = \{p^b, p^c\}$$
$$O_3 = \{p^a, p^c\}$$

Then we can check that $\{\{O_1, O_2, O_3\}\ ;\ \{p_a, p_b\}\} \looparrowright \{\{O_1, O_2, O_3\}\ ;\ \{p_a, p_b, p_c\}\}$.
Or, as a slightly more complicated example:

$$O_1 = \{p^a, p^b, p^1, p^2\}$$
$$O_2 = \{p^c, p^2\}$$
$$O_3 = \{p^a, p^1\}$$

Then we can check that $\{\{O_1, O_2, O_3\}\ ;\ \{p_a \cup p_1, p_c \cup p_2\}\} \looparrowright \{\{O_1, O_2, O_3\}\ ;\ \{p_a, p_b, p_c, p_1, p_2\}\}$, recalling (6.17) above.

[Sup2] Pure serendipitous covariation The properties in Π_i and Π_j reduce to sets that have no fundamental properties in common, so that $\forall P \in \Pi_i, \forall Q \in \Pi_j : P \cap Q = \emptyset$ but each property in the reduction of Π_j is accompanied, serendipitously, in objects by one or more properties from the reduction of Π_i

[Sup3] Pure supervenience The properties in Π_i and Π_j reduce to sets that have no properties in common, so that $\forall P \in \Pi_i, Q \in \Pi_j : P \cap Q = \emptyset$ but each property in the reduction of Π_j is always accompanied in objects, for some ontic reason, by one or more specific properties from the reduction of Π_i

We can, of course, produce hybrid cases by combining aspects of all three versions: they suffice to build up most, if not all, possible modes of supervenience.

6.3.7 Pure supervenience and causality

What is now useful to do is to investigate the causal relationships between the properties in **[Sup3]** as we need this in our discussion of the effects of supervenience on the ideas in this study. It is possible to look at the causal relationships for **[Sup1]** and **[Sup2]** as well but we shall not do so as they distract from our central purpose.

To simplify matters it is better to base our discussion around a simple example of pure supervenience rather than seeking generality at this stage. We shall start with a toy universe populated by a few objects with a very limited set of properties. As before, we label an object by the properties it possesses, so an object possessing p^1 and p^2 would be written as $\{p^1, p^2\}$ and, according to equation (6.3), we have:[39]

$$p_1(\{p^1, p^2\}) = 1 \tag{6.63}$$

We set out the objects in our toy universe as follows:

$$\begin{aligned} O_1 &= \{p^a, p^1\} \\ O_2 &= \{p^b, p^1, p^2\} \\ O_3 &= \{p^c, p^2\} \end{aligned}$$

Now consider the two pre-topologies:[40]

$$\boldsymbol{T}_1 = \{\{O_1, O_2, O_3\}\ ;\ \{p_1, p_2\}\} \tag{6.64}$$

and

$$\boldsymbol{T}_A = \{\{O_1, O_2, O_3\}\ ;\ \{p_a, p_b, p_c\}\} \tag{6.65}$$

It is easy to check that $\boldsymbol{T}_1 \looparrowleft\!\!\!\rightarrow \boldsymbol{T}_A$ and we can explain this as being either an example of **[Sup2]** or one of **[Sup3]**. It cannot be an example of **[Sup1]** as we have used fundamental properties in our universe and these do not reduce. We shall only consider the case of **[Sup3]** although our discussion applies in the most part to an analysis of **[Sup2]** as well.

Pure supervenience here says that in our universe whenever an object has [an instance of...] property p_a it must also posses property p_1 and it further says there is an ontic link between these properties - an [ontic] supervenient relationship between them. Similarly, whenever p_b occurs in an object, it must be accompanied by the properties p_1 and p_2. There is an [ontic] supervenient

[39] We assume these objects possess these properties at all times.

[40] In light of our earlier discussions, we do not worry about Boolean completion and using topologies instead of pre-topologies.

relation between p_b and the two p_1 and p_2. And similarity between p_c and p_2. So it appears that occurrences of properties in the base set necessitate, in some manner, the occurrences of some of the properties in the supervening set. The reverse is not the case as we can see the occurrence of p_2 does not necessitate the occurrence of any particular property from the set $\{p_a, p_b, p_c\}$. We shall write these ontic supervenient necessitations, which we shall call 'emergence', as follows: [41]

$$\{p_a\} \Rsh \{p_1\} \tag{6.66}$$

$$\{p_b\} \Rsh \{p_1, p_2\} \tag{6.67}$$

and

$$\{p_c\} \Rsh \{p_2\} \tag{6.68}$$

Where we take these to summarise a general relationship between property instances.

And we take it that emergence is the underlying *ontic* relationship that gives rise to supervenience between sets of properties. And, further, we note that emergence is an ontic relation between properties occurring *at the same time* and that, in contrast, causality is a relationship between proprieties occurring at differing times.

We now turn to examine how the emergence relationship fits in with causal equivalence from earlier. The point to note is simply that the two relationships, causality and emergence, are entirely distinct relationships. Nevertheless we could require that the emergence involves causal equivalence in some form but only in universes that permit causal over-determination. So that for sets of properties $\{p_1\}$ and $\{p_a\}$, we might have:

$$\{p_1\} \Rsh \{p_a\} \Rightarrow p_1 \triangleright p_a \tag{6.69}$$

And this may then lead to more general relationships between sets of properties such that, when $\Pi_2 \looparrowright \Pi_1$ as pure supervenience, we might have a relationship of the form:[42]

$$\Pi_1 \Rsh \Pi_2 \Rightarrow \Pi_1 \triangleright \Pi_2 \tag{6.70}$$

[41] It is interesting to compare this with the idea of emergence in Bedau (1997) and Bedau & Humphreys (2008) and Butterfield (2011). Butterfield takes the view that supervenience is a relationship between theories and so adopts a broadly Nagelian, see Nagel (1961), view of reduction in his work.

[42] We do not define the terms precisely here.

But it is important to note that there is nothing about pure supervenience itself that requires this relationship.

In a more detailed analysis of the relationship between pure supervenience and causality it would be fruitful to investigate different ways that the causal equivalence of property sets could be unpacked. Further, whilst we stated above the causal and pure-supervenience relationships need not be related, it must be the case that there are some joint restrictions that arise between them. It may be the case that the essential nature of the universe, via the ontic laws, is set in such a way that it respects the pure-supervenience relationships; or it might be that pure-supervenience is actually no more than a facet of a certain restriction on the ontic laws of nature. However, it would seem that the universe must conspire to ensure that the correct group of properties are caused together in an object if they are to supervene. However (6.70) would seem to be far too restrictive to be the relationship that links causality and pure-supervenience. Much more work is needed here.

It is useful to note that we cannot require both (6.70) and pure supervenience in the absence of causal over-determination. If we ban causal over-determination then the right-hand-side of (6.70) would only be the result of a straightforward reduction of the properties as in **[Sup1]**. It would be interesting to examine Kim's argument for physicalism, as set out most recently in Kim (2007), using the techniques here. We leave that, however, for another time.

6.4 Does supervenience affect our causal arguments?

In this final section, we marshal the work of this chapter and use it to advance a case for a claim. The claim is that our previous arguments are unaffected by the possibility of supervenience. In simple terms, our work above shows that, in most cases, supervenience is not an additional ontic relationship so its presence does not affect our arguments based on causality.

There is an exception to our claim, albeit an exception that need not worry us unduly here. We must consider what pure supervenience might bring to the mix. Indeed the exception is that there might be a form of dualism between the mental and the physical mediated by pure supervenience. This possibility might be a route to making the claim that the mental and the physical are entirely separate realms. Recall, at the end of section 5.5 we suggested that if we were to make this claim then we would have to deal with the difficulty of how we can justify asserting that our mental has access to the physical. The answer may well reside in postulating pure supervenience as an interface between them. However, here we *choose* not to countenance dualism because it requires an extension of the idea of a 'universe' beyond the definition of section 3.7. Nevertheless, the possibility is an interesting one and suggests further work may be fruitful.

First, then, the reason that our arguments are mostly unaffected by supervenience is that in all but one case of supervenience, the exception being pure supervenience **[Sup3]**, the relationship is merely one that derives from either serendipity or property reduction. As such they add no new *ontic* relationships beyond the causal ones we have used throughout and our arguments are unaffected.

The one case that may well have some implications for our arguments is that of pure supervenience and it is to examine this we now turn.

This is our situation: we have a universe of objects and we find that some of the properties of the objects divide into sets that are superveniently related. We then claim that it is possible that the relationship is more than serendipitous and derives from some underlying ontic nature. We claimed above that supervenience arises from a more basic relationship between properties that inhere in objects, namely emergence. We ask if this additional ontic relationship between property sets that cluster together around [or make] objects has any implications on the arguments we have presented in this study.

There are two places where additional supervenient relationships might affect our arguments one is in the assumption we made concerning restrictions when one relatum has access to another. Recall we claimed in equation (5.1) of section 5.2 that if relatum A is to have access to some prior relatum B then we must have:

$$\mu(B, A) = \mu(B)$$

However, the content of a relatum is of a finite size and our contention that if one relatum is to have access to the form of the content of another relatum then it must obey (5.1) regardless of whether the relata are contemporaneous or not.

And the second place is as an additional option to our discussions following Theorem 1.

Before we examine how supervenience might affect our arguments, we take a technical detour: we first need to clarify whether the property sets we are dealing with arise within a single universe or straddle two universes. The reason for this is that previously we defined, in section 3.7, universes using purely causal ontic links but now we have the possibility of a second form of ontic link, namely supervenience. It may now be the case that distinct universes that are each *causally* closed can nevertheless be joined together by these supervenient links. This calls into doubt the way we set about defining what we meant by ‘universe’.

If pure supervenient links can only arise within a single universe, causally construed, then our ontic supervenience relationship, and thence emergence, need only be considered as an ontic addition within a universe. If supervenience can straddle universes, causally construed, as an ontic relationship then we find that the causal mechanisms we used to delineate a universe only close that universe to outside *causal* influences but not to supervenient ones. Informally, the possibility that we have two causal realms linked by a supervenience relation can be thought of roughly as being like two worlds persisting in parallel.

If we consider a case of **[Sup3]** for $\{\mathbb{O}, \Pi_i\} \looparrowright \{\mathbb{O}, \Pi_j\}$. If we also recall the definition of a causal nexus generated by any relatum that we gave in section 3.7, we can then consider the causal nexus generated by any relatum corresponding to any instance of any general property in Π_i, denoting the nexus as $\mathfrak{U}_i(p)$ and similarly for Π_j, denoting the corresponding nexus by $\mathfrak{U}_i(q)$, where p and q are relata associated with the respective property sets. There are three possibilities for these causal nexuses: one is that they are in fact the same nexus and one is that they are entirely separate causal nexuses. The third is a that they are some mixture of the two - we do not consider this case as it is covered by our arguments for the first two cases. That is, we have either:

$$\mathfrak{U}_i(p) \equiv \mathfrak{U}_j(q) \tag{6.71}$$

or

$$\mathfrak{U}_i(p) \cap \mathfrak{U}_j(q) = \emptyset \tag{6.72}$$

We can interpret these as saying that, in the case of (6.71) that both sets of properties inhabit the same universe and, in the case of (6.72) they form causally distinct universes. In the case where they are in fact the same causal nexus then all we have done is discovered a synchronous ontic relationship between relata and this does not affect the arguments we have advanced earlier.

First, we offer some brief speculations concerning the dual case: In this case, where the two sets of properties inhabit causally distinct realms, and hence causally distinct universes, we have a more unusual scenario. If (6.72) is true then we have, in effect, two [causally closed] universes that evolve in such a way that some aspects of one universe reflect some aspects of the other universe even though they have no *causal* effect on each other. If we use the terms ‘mental’ and ‘physical’ - one for each realm - then this means that we have a mental realm and a physical realm which are causally independent and yet have some non-causal relationship through supervenience.[43]

We shall now argue that accepting two causal realms and maintaining the ontic nature of the supervenience relationship must lead to a radical reappraisal of the idea of an object and difficulties that are hard to resolve without further work. If we return to the way we defined supervenience, we see that we always had supervenience with respect to sets of objects. Objects are the way we take properties to be grouped, they tell us when certain properties occur together. Then, in section 6.3.2, we defined objects as clusters of fundamental relata, or as just relata, and held that not any cluster could be an object but rather objects must be in some sense 'natural'. We left it vague as to whether specifying properties, as causal aspects of relata, exhausted the relata or merely referred to some aspects of the form of their content in its causal capacity. That is, we left it vague as to whether objects are mere clusters of properties or whether properties cluster around, or perhaps inhere in, something else to make objects. In our dual system we then have two views we can pursue for objects: they can be mere clusters of properties alone or they can be underpinned by some additional 'substance', where we use the term to indicate the ontic underpinning around which properties cluster. If we accept an underlying substance then it must, in the dual case, straddle two causally isolated universes. What is not clear, in this case, is what it is to be a substance and to exist in two causally isolated realms. That is not to say the possibility is incoherent and can be dismissed; rather it is to say that the ideas need more work if they are to have plausibility. The other case, that objects are clusters of properties, is perhaps a better route to take here. We said above that the clusters must be natural ones, perhaps reflecting the locality we see in our everyday conception of objects, and this suggests that supervenience is a relationship that must then be part of the defining nature of objects *qua* clusters of properties. If we accept this and leave the exact nature of the clusters within a universe vague then at least we have some purchase on trans-world supervenience.

In this case we have a tentative model for universe pairings: we have two isolated causal nexuses with clusters of properties in one universe - perhaps 'sub-objects' - connected with 'sub-objects' in the other universe. And the connection is mediated by an ontic supervenience relationship.

We shall say very little more on this. It is not clear what to make of the proposal and it needs much more rigorous underpinning and exploration if it is to be proposed as a plausible model of pure supervenience and dualism. What is relevant here is to ask whether such a model destroys any of the arguments we have advanced in this study. In a certain sense it does as it opens the possibility that the relata in each universe are of a fundamentally different nature and so opens up the possibility of talking of 'mental substance' and 'physical substance' as being fundamentally distinct. If

[43] We use the terms 'mental' and 'physical' loosely here.

we grant this, we must ask a second question: we must ask what is it that gives the claim that the mental is in some manner structured so that it has access to the machinations of the physical. That is, how do we know what we have some access to the physical world if we are dual mental beings - an old question. Previously we claimed that this access could be justified using causal relations but we do not have this option here. Instead we must claim that it must be justified by appealing to the supervenience relationship between the property sets in each universe and hence parasitic on the emergence relationships. However, it is not clear what to make of that as a suggestion and nor is it clear how to unpack the concepts rigorously to make them plausible. Instead we shall leave the possibility of parallel universes 'glued' together by supervenience relations for another time. Nevertheless, it is worth noting this possibility does not affect Theorem 1 directly as we set it as an assumption but it does add a small counter to one aspect of the discussion we offered. We mentioned that if we were to take the mental and the physical as two realms then we must also seek to provide some explanation as to how the mental gains access to the physical. Here we must concede that pure supervenience might be used to construct this interface. More work needs to be done to clarify the scope of this tentative suggestion

Returning to the question as to whether supervenience as a synchronous ontic relationship *within* a universe affects our arguments. We note that the supervenient relationship might well alter the dynamic between the content of relata at different times and might even allow contemporaneous relata to have access to each other in some manner. However, we argued above that (5.1) is unaffected and further none of our arguments hinge on the absence of extra restrictions on the way that relata occur moment to moment or the way they interrelate at each moment. As such we hold that the possibility of synchronous ontic relation of supervenience does not affect our arguments if they are based within a single causal nexus.

So whilst supervenience provides some interesting and intriguing possible additions to the relationships that may exist within a single universe, as a closed causal nexus, it does not affect our single-world causal arguments.

In the absence of pure supervenience across universe pairs, our arguments stand in the face of pure supervenience.

Chapter 7

Conclusion

7.1 Introduction

In this, the final Chapter, we provide a critical summary. In the first part we revisit the mechanism for the exercise of free will that we proposed; and in the second part we critically appraise our metaphysical models of the universe, causality and supervenience.

7.2 PART I: The central argument and its limits

We have argued that our ontic model for free will - of the locally-deterministic self and of immanence combined - fulfils the two criteria we set as being necessary for free will in Chapter 1. Recall, we set these necessary conditions as:

> **1.** The universe presents genuine choices, genuine alternative futures, to its inhabitants;
> **2.** The inhabitants 'choose' between these choices in such a manner that they are clearly and solely responsible for their choices, their selections.

First we shall explain how our arguments mesh and then we shall look at some areas where difficulty might be seen to remain.

If we were to derive the idea of choices as the presentation of alternative futures directly from indeterminism in some form then, by requiring **1**, we would implicitly assume that no compatibilist solution to the problem of free will could launch. What we did, however, was to show that such an idea of free will cannot be eked out from determinism alone and so concluded that **1** and **2** are antithetical to compatibilist stances if genuine ontic choices and sole responsibility are required.

In Chapter 4 we argued that ontic choices, the possibility that the universe could go one way or the other at some point, can only be cashed out in terms of immanence. And in doing so we drew heavily on the ideas we set out in Chapters 2 and 3. We showed that determinism and choice, even crudely construed, cannot mix if ontic restrictions are to be respected. We put this argument both using the detailed work we presented on causality in Chapter 3 and using the incoherence of universe-separability that we argued for, and proved, at the end of Chapter 3. That is, we took choice away from the conceptual sphere and placed it firmly in the ontic sphere. And in doing so we discovered that indeterminism of a specific sort, namely immanence, is the key to an ontic response to **1** and **2**.

Once we had an ontic notion of choice, as immanence, we required a notion of the choosing, the very act where one choice is selected over another. And we required this choosing to be such that it clearly warranted the attribution of responsibility to the choosing agent. However, in introducing the idea of choosing between presented choices, we were constrained by Theorems 2 and 3 of Chapter 5. Recall, these showed that piloting was not an option. Of course, it is true that in all cases of choosing there must be an interaction between the choices on offer and the choosing agent. However, it is usually the case that this process is taken to be accomplished in stages: we usually take it that choices are presented in their raw form, as clear alternative futures, to an agent who is aware they are choices,[1] and that they are presented as genuine choices - whatever that means. The agent then takes time to pass these possibilities through their thoughts and to deliberate actively towards a selection. And this active deliberation is taken as being a deliberation where the self, the choosing agent, is somehow in control of the process - a pilot guiding the ship of consciousness. But recall, we rejected this thinking as we proved, in Theorems 2 and 3 of Chapter 5, that deliberation cannot be piloted; instead it must be just a natural evolutionary process in the universe. We rejected the multi-stage process choices-to-choosing, namely awareness-deliberation-selection. In place of deliberation as the factor that reflects the self's responsibility for their choice, we put the notion that the self must evolve in a manner consistent with its nature at each moment - so that its nature is both a reflection of what it is and responsible for its future nature. And we said that this can only happen if the self evolves in a manner that is deterministic in the main. That is, we replaced the idea that deliberation-to-choosing is the mark of a responsible self with the idea that consistency - as being unavoidably true to ones nature - is the mark of responsibility. We then used these two facets of our model - presented choices through immanence and responsibility through being true to, or consistent with, one's own nature - to propose that the exercise of free will, as set by our criteria, comes in their combination. And we argued that to keep responsibility the result of their interaction must be determined and not open to counterfactual truth. That is, in rejecting de-

[1] We make no issue of the moves between the world and our concepts of the world here. We assume that, in this part of the discussion, the presentation of choices is wrapped up suitably with the agent's being able to recognise them as choices - whatever the choices might actually be

liberation as the mark of the self's responsibility, we rejected the counterfactual encapsulation of responsibility: we rejected as incoherent the idea that responsibility must rest on the maxim 'could have done otherwise'. And we explained how sole responsibility derived from the serendipitous encounter between an immanent relatum and the self.

Our model is certainly not standard. It admits that an [ontic] exercise of free will that satisfies **1** and **2** above is possible but achieves it in a manner that is not found elsewhere in the literature. First, we can say that our model would fall broadly into the incompatibilist camp. The arguments we advanced in Chapter 5 show indeterminism and immanence are essential to our mechanism. However, our solution is not in the mode of standard incompatibilist solutions either. Rather, it combines both a local determinism, in the evolution of the self, and a modified form of indeterminism to allow for immanence. Nor can our solution be accused of being a Valerian-type one. Valerian-type solutions, recall, keep hold of the idea of deliberation as a motion from choice to choosing but inject uncertainty, through indeterminism, somewhere into the very deliberative process as a proxy for choices. We do not do that: we reject both the part played by deliberation-to-choice, as a result of Theorems 2 and 3, and we reject the possibility that indeterminism in the selection itself can lead to responsibility. As such we claim our solution is novel both in the way it cashes out responsibility as local-determinism and in the way it introduces choices through an ontic idea of immanence.

In brief then, our rejection of the part that deliberation plays in choosing led us to make local-determinism the mark of responsibility. And in a search for genuine choices, as genuine alternative futures, we rejected determinism and pure indeterminism; instead we introduced immanence as a facet of indeterminism. We then sought to combine them in such a manner that maintained sole responsibility through local determinism in the face of the indeterministic nature of immanence and succeeded in producing a model that satisfies both **1** and **2**.

This all being said, our model has many shortcomings and elements that are less than satisfactory as they stand. But nevertheless, we commend it as a first tentative step in looking at the exercise of free will from a new and purely ontic angle.

We turn, now to look critically at our free will model and leave our examination of the metaphysical models we have met along the way until the following section.

There are three major aspects of our model and arguments that might be considered as shortcomings for varying reasons. All three stem from the sterile nature of our model - the sterility that traces back to our 'ontic causal monism' of Theorem 1: of leaving aside issues of mental lives, thoughts, feelings and the like.

The everyday notion that first we become aware of choices, alternatives, and only then deliberate prior to selecting between them is not easy to reconcile with our alternative mechanism. Usually, we take it that choices are presented to us and then, and only then, do we move from an awareness of choices to choosing. Our model allows no moment for an awareness of choices to play a role; rather, we have proposed a sterile combination of immanence and self that immediately leads to a deterministic sequence. That is we must claim that awareness-of-choices is not a significant part of the exercise of free will as we have it.

We must then pose two questions: one is to ask whether our notion of choices as immanence is the origin of the illusory awareness of genuine choices that we often have: I am presented with apple pie and cheesecake and I feel, when they are presented to me, that there is yet a decision to be made. If our model is correct, it may be the case that the presentation of the two desserts is framed as an immanent relatum but that the moment of interaction with the self leads to the illusion that a genuine choices [whatever that is in this context] still obtain whilst, in reality, the path that my deliberation will take is set by the very mechanistic details of the self at that moment. Second, we must also ask, if it is the origin of the illusion how is it so and why is it so? To this, we have no answers to offer. And so we leave the issue of the relationship between immanence and the perception of choice as an open question.

Second. We have argued for, and proved for our model, that deliberation-to-choice is not a piloted mental act. At most, it is merely a manifestation of the machinations of the self as it moves its path from having made a choice to awareness-of-choice and then, perhaps, to action - the eating of pie. This, at first glance, seems rather an uncomfortable position. We certainly *feel* that our deliberation are not mere mechanical processes but are a facet of our making choices and resolving what is unresolved. That is, we take it that our thought processes lead *to* selections. This conflict between our intuitive feelings and our version of free will is an uncomfortable aspect of our model. It is not easy to become comfortable with. We can say that, perhaps, the human mind has evolved to react to the initial combination of self and immanence in a manner that gives the consistent illusion that deliberation leads to choice. And the consistency of the illusion would have to be quite sophisticated.

In fact the deliberation may be just the 'mental' aspects of the machinations of the self in moving from choice to eventual action. Why we should have evolved this way, why we need mental lives and why they should have some kind of inner consistency to them to perpetuate this illusion is difficult to unpack. At best we must either accept our model as correct and then move to consider why we have mental lives at all, and why we perceive deliberation to be important; or, if we are not bothered by its importance, we must explain its presence.

However, the process of deliberation does *seem* important. If we permit a dog, as a convenient example, to be a self and admit that dogs are lesser creatures that lack such attributes as self-awareness, then we must allow dogs to both exercise free will and be responsible for their actions under our scheme.

If we wish to eliminate dogs and so forth from the realm of responsible agents, it does then seem that deliberation must be important. Perhaps what is needed to bridge some of this gap, at least to eliminate dogs and the like from the mix, is a move from responsibility to moral responsibility. Perhaps we must require that a self is not only responsible but must possess a complex internal mental life with some specific attributes - such as self-awareness and a *sense* that they are responsible for their choices. And this may manifest, in part, through *believing* piloted deliberation leads to a selection. Why we should have evolved this way, however, is not so clear. Perhaps there is some advantage to developing a sophisticated idea of moral responsibility that explains the presence of a sophisticated consciousness in humans. We leave this as an open and unresolved question here.

Third. When we set out our version of the exercise of free will we mentioned that it may well be possible that there are many ‘objects’ in the universe that satisfy the defining characteristics of a self. If these ‘objects’, on combination with an immanent relatum, lead to a deterministic outcome then they must be held responsible for the outcome. As this appears to allow for the possibility of objects that possess no consciousness having responsibility attributed to them, we added the speculation that the only way a self can combine with an immanent relatum to give a deterministic outcome would be if the self possesses specific extra aspects - such as a consciousness in some form. That is, we attributed to the presence of consciousness the power to play a crucial part in our mechanism. This was an *ad hoc* speculative solution to a problem that is not easy to resolve. Either we must accept that our model allows for some inanimate objects to exercise free will, or we must add very specific ontic properties to consciousness, or we must take pause.

Pausing is perhaps the best option here. It would be plausible to claim that our notion of the exercise of free will is incomplete as it stands. This is in keeping with what we hinted at in Chapter 1. There we stated that we sought criteria **1** and **2** for our model and that we took these to be necessary conditions. We did not, however claim that they are the only necessary conditions. If there are more necessary conditions then we should ask some further questions:

> What are the additional necessary conditions?
> What is the rationale behind the additional necessary conditions?
> Can our model accommodate the additional conditions or do they clash with our model and show that free will is dead in the water?

Until these are resolved we cannot make the claim that our model is complete as a resolution of the issue of free will *in toto* even though our claim that, as it stands, it is ontically optimal as a model satisfying **1** and **2** is unaffected. If further necessary conditions do indeed clash with our ideas, we would be forced to claim that, if we want **1** and **2** to stand, free will is defeated and then move to agree with those who claim it is nothing more than an illusion.

What these three issues amount to is this: we need either to move our model forwards to accommodate more than it presently does or we must admit that the uncomfortable sterile picture it leaves us with is one that is both correct and complete. If we choose to move, then we must either seek to challenge **1** and **2** or seek a way of moving to look inside the relata, to see how the mental aspects fit with our picture of free will. The arguments we have presented stand as a strong defence of **1** and **2** so we must, if we will not allow sterility, take the second option.

We must probe deeper to expunge the uncomfortable feeling that our model has left lingering. And we must do so whilst remaining cognizant of the spilt between the ontic and the epistemic, the Archimedean view and the internal view. We must be aware that we may well have exhausted the purely ontic necessary conditions for free will and that our criteria, **1** and **2**, are optimal as ontic conditions. Perhaps, then, what remains are necessary epistemic conditions - they are what is required whilst looking out from the inside rather than what is required when looking in from the outside. And then, if our additional necessary conditions are epistemic we must seek to justify why this is the case - and perhaps some of the issues we adumbrated above provide the initial impetus. If we cannot resolve these issues then we must admit our model is strong up to its boundaries but begins to crumble when we push those boundaries a little further. Whether our pushing is legitimate is left as an open question.

7.3 PART II: Our metaphysical models

7.3.1 Introduction

In this section we review some aspects of the models we have used throughout this study and suggest where future work might be fruitful.
There are three broad issues we shall consider here when reviewing our models:

A. How realistic are they as models of the world?;

B. What modifications and extensions to our models might be useful?;

C. What claims to originality can we make for our models?

Below, we shall discuss each of these issues in turn. It might seem that moving towards providing answers to **A** and **B** is important and we would not disagree. However, we must be careful to clarify where such answers are important and where they are not. Recall, two of the aims for this study: to provide a model for the exercise of free will that satisfied some criteria that we set out in Chapter 1; and to provide a set of mathematically-couched metaphysically inclined models for causality, determinism and supervenience. For the first aim, all we strictly require is *some* model of a universe that does allow us to demonstrate a version of the exercise of free will that satisfies our criteria. And we claimed in Part I above that the model we have offered fulfils this. Of course, it would be desirable if the models we have offered in pursuit of our first aim are plausible but this is not a strictly necessary condition of fulfilment. For our second aim, both **A** and **B** are much more relevant as question and, additionally, some answers to them would add weight to the free will model we developed.

What the discussions below amount to is: first, that our models are robust and consistent as far as they go; second, they leave scope for extension into more sophisticated structures; and third, that issues of plausibility would benefit from a careful and extensive comparison with standard physical and philosophical models.

The strength of the models as we have presented them is that they have given us sufficient scope to investigate a sterile version of free will; and the weakness of our approach is that we do not provided substantial comparisons with other models of causality and determinism. We take the latter as a lacuna of the study - resulting from limited space - rather than a flaw in the models themselves.

We now turn to answer **A**, **B** and **C**.

7.3.2 How plausible are they as models of the world?

In answering **A** in relation to our models *qua* models, as opposed to our use of them in our version of free will, we can break down the questions a little further:

A1. Are our models plausible models of the world judged from:

a. Our everyday experiences;

b. Current physical theories and experiment?

A2. Do our models obey the Archimedean constraint?

In part answer to **b** we offer the discussion and comments we have made throughout the main body of the study. For instance, we have commented on:

- The issue of discrete time and the possibility of modifying our models to accommodate continuous time;
- The failure to take account of special relativity and other views on time;
- The nature of the connexion - whether it is a flow of information, energy-momentum or something altogether different;
- The relationship between relata, causality and an underlying spatial manifold;
- The conservation of measured quantities;
- The nature of objects and their persistence through time;
- The nature of properties and the lack of structured relationship for them within our models;
- The clash between our notion of property and that appearing in some interpretation of quantum theory.

And we would claim that the models we have offered do not stand in any substantial opposition to any current physical theories or experiments but we must leave that as a tentative claim open to challenge.

There are, however, two areas in answer to **b** that do merit some brief further comment: one is the way we have dealt with time; and the other is the way we have required all our structures to be discrete and either finite or countably infinite.

First, time: we have taken a naïve view of time. As we discussed in Chapter 2, we take time to be discrete and defined in such a manner that there is a clear globally delineated present. This could be taken to stand in opposition to special relativity where time is taken relative to inertial reference frames.[2] This does then suggest that our model lacks plausibility with respect to its stance on time as an absolute - leaving aside the issue of discreteness for the moment - and this is an accusation that needs addressing. We can offer two answers: we can admit that our model is lacking and work to produce modifications that are sympathetic to a relativistic view of time; or we could attempt to defend our naïve model of time and point out that experimental verifications of special relativity always involve duration related to change between two spatio-temporally separated points and it is possible to claim that our model would be able to incorporate various modifications of the rate of processes, such as rate the decay of particles, as it stands and so furnish potentially the same sorts of theoretically predicted results as special relativity. This second answer is, however, a bold claim and certainly more work is needed. It is also somewhat tentative - especially when considered together with our comments in Chapter 2 on the absence of duration in discrete time models.

Discrete versus continuous time is equally tricky and we have already offered some comments on this in section 2.4.

The other issue we mentioned concerned the need for finite or countably infinite aspect to our models: we used this when dealing with the combination of fundamental relata and the related issues of Boolean completion of property sets, as well as when we set out our measure and claimed the measure applied to the whole universe gave a finite answer. Certainly our motivation in these cases derived from a desire to use certain mathematics to articulate our thinking rather than any underlying metaphysical issues. However, if we were to abandon countability then it is not easy to understand how alternative models could be built. Once again, more work and clarification of this point is needed.

We now turn to answering **a**. This is perhaps somewhat more tricky as our everyday experiences do not seem to suggest that we are enveloped by, or formed from, fundamental relata or that we can assert the conservation of some measurable quantities and so forth. In response, we make two comments: first, as our discussion concerning **b** suggest, the models we have offered are plausible models for the underpinnings of the everyday in the same sort of way that quantum theory is a plausible model for the everyday - in as much as it gives correct predictions to experiments and so does not wildly *contradict* everyday experience. And second, we would suggest that our everyday experiences of the world, of cars and cattle and the like, are not good guides to what constitute plausible models. We would claim that the only guides we should take to the plausibility of models are the scientific models, somewhat loosely construed, we currently have and the experimental data that accompany them. And recall, that we set limits on what they can tell us in Chapter 2 when we

[2] For a more precise and mathematical treatment see, for instance, Schutz (1985, Chapter 1) and for a more philosophical treatment see, for instance, Bohm (1996).

differentiated predictability and determinism.

What this amount to is that the strength of our answer to **A** must rely on our answer to **b** and that our answer to **b** only ever allows us to claim our models are plausible, in the absence of contradiction, and does not allow us to assert they are correct.

Now to **A2**. We have placed restrictions on the way we, as denizens of a universe, can then access our habitat. We placed constraints when we set out **[EA]** - as discussed in sections 2.8 - and further technical restrictions when setting the assumptions for the three theorems in section 5.2. Recall that the Archimedean Constraint, discussed in section 2.5, is one of consistency. We do not need to demonstrate that our restrictions still allow us to make the metaphysical claims encapsulated in our models but rather we must demonstrate that our models do not stand in contradiction, or potential contradiction, to the making of such claims. We hold that there is nothing in our models that tells us we cannot make such claims; as such, the constraint is obeyed.

7.3.3 What modifications and extensions to our models might be useful?

Here we concentrate on three issues that we see as underdeveloped in this study and worthy of further study: the first is the relationship between the causal structure of the relata and the manifold structure of the world around us; the second is the need for a structured set approach to supervenience; the third is the relationship between counterfactual notions of causality and our connexion-bound approach.

Throughout this study we have attempted to separate out the idea that relata are manifold bound. Recall, we suggested this as we did not want to engage with issues concerning a manifold and the causal or ontic nature of 'raw space'. Additionally implicit within our rejection of a distinct underlying spatial manifold were some further motivations which we hinted at: that there are reasons - stemming from issues in physics - to abandon the manifold and inherited notions of locality as primary.[3] Two areas of further development would be useful here: one is to undertake a careful philosophical study of spatial notions and locality and how they constrain and arise in metaphysical models;[4] and second, to develop our models to incorporate spatial notions into our work whilst taking account of both metaphysical investigations of space and of physical theories that engage directly with spatial, and spatio-temporal, ideas.

Now we turn to the possibility of developing a structured sets version of supervenience. We have already made a number of comments on this in the second half of Chapter 6. Here we concen-

[3]Our prime motivations stem from a certain attitude toward difficulties in quantum theory - see, for instance, Callender & Huggett (2001). Additionally there are some interesting discussions can be found in Seiberg (2005) and Dowker (2006).

[4]Some useful work on locality and quantum theory can be found in Redhead (1989).

trate on making one further comment that we hold shows there is a need both to extend simple supervenience to structured sets and to develop a notion of structured sets in a spatial context. The issue that we take motivates this is the use of supervenience in relation to the mind-body problem. It is often held that the mental supervenes on the physical. We presented a crude flavour of this in section 6.2 where we gave an informal overview of supervenience. However, what seems to be claimed in such analyses is that identical [physical] brain states give rise to identical mental states.[5] But identity of brain states requires, again in simple terms, that each brain be in a complex state requiring neurons to be placed in certain relative positions, for neurons to be firing or not firing in certain ways and so forth. In other words a full property specification of a brain [in 'physical' terms] requires both properties *and* [relative] positions to come into play. As such any property description of a brain state would have to involve some spatial structure.

There would seem to be much interesting and potentially fruitful work and clarification that might arise from such an approach.

And third, it would be interesting and fruitful to investigate the relationships between our models of causality, over-determination and immanence and those approaches that rely on counterfactual encapsulations of causality. We could ask how we might set about constructing a systematic comparison of the approaches and then ask if a careful analysis using these three strands - causality, immanence and over-determination - leads to the same broad results in both approaches. And if it does not lead to the same results, what can we learn?

7.3.4 What claims to originality can we make for our models?

Originality is hard to pin down precisely as shadows of others' work and ideas fall across and influence our work here. There are two areas where we might consider the originality of our models: one is the *use* to which we put them; and the other is the details of the models themselves. We have dealt with the use we envisaged for our models in Part I above. And there we held that our models have succeeded in demonstrating that a certain sterile version of free will, satisfying the criteria we set ourselves, can be constructed. As such, here we concentrate on making some brief comments about the originality of the structure and ideas behind our models,

In all three cases, we cannot claim any originality for the general motivations behind our models. Connexion-bound universes are found elsewhere in the literature; causality based on the idea of transference is found elsewhere in the literature; and supervenience is found elsewhere in the literature.[6] That is, the overarching notions we use are found elsewhere. However, in all three cases

[5] This is a slight over-simplification but it will suffice for the point we wish to make.

[6] We refer the reader to the literature references in the respective Chapters.

we hold that the methodology of our approach and many of the conclusion we have deduced are original. And we hold that the use we have put our models to in this study demonstrates they have some claim to being philosophically useful.

As an exercise, we list what we consider to be the main original results and assertions that we have derived or made using our models. Some are dealt with extensively in our study and a few merely glanced at in passing:

- Determinism is an ontic doctrine based on connexions;
- Determinism and regularity do not necessarily relate;
- We cannot tell what sort of universe - deterministic or other - we inhabit;
- Causality and determinism come apart;
- Determinism requires global essentialism but permits local separability;
- No individual can pilot themselves - so deliberation to choice is dead in the water;
- Ontic counterfactual analyses of choice require immanence;
- Choices, as ontic future possibilities, can only arise from immanence;
- Responsibility can arise only from local determinism;
- The ability to do otherwise cannot be the mark of free will;
- We cannot know if we have free will;
- Pure supervenience arises as an ontic synchronous relationship.
- A form of emergence might underpin supervenience.

And these stand as an assertion that, in some small part, aspects of the study presented here have a claim to originality.

7.4 Envoi

At the start of this study we set out three broad aims: to produce a version of free will satisfying two specific criteria; to develop models for causality, supervenience and determinisms; and to set a foundation for future work.

Above we have put the case that we have achieved all three of our aims.

As ever, there is much left to do . . .

Bibliography

Armstrong, D. M. (1985). *What is a Law of Nature? (Cambridge Studies in Philosophy)*. Cambridge University Press. 44

Atmanspacher, H., & Bishop, R. C. (2002). *Between Chance and Choice: Interdisciplinary Perspectives on Determinism*. Thorverton UK: Imprint Academic. 20

Barbour, J. (1999). *The End of Time. The Next Revolution in Our Understanding of the Universe*. Weidenfeld and Nicolson. 35

Barrow, J. D., Tipler, F. J., & Wheeler, J. A. (1988). *The Anthropic Cosmological Principle (Oxford Paperbacks)*. Oxford University Press, USA. 34, 38

Bedau, M. A. (1997). Weak emergence. *Philosophical Perspectives*, *11*, 375–399. 188

Bedau, M. A., & Humphreys, P. (2008). *Emergence: Contemporary Readings in Philosophy and Science (Bradford Books)*. MIT Press. 188

Beebee, H. (2004). Causing and nothingness. In J. Collins, N. Hall, & L. Paul (Eds.) *Causation and Counterfactuals (Representation and Mind Series)*, (pp. 291–308). MIT Press. 63

Beebee, H. (2006). Does anything hold the universe together? *Synthese*, *149*(3), 509–533. 23

Beebee, H. (2011). *Hume on Causation*. Routledge. 23

Beebee, H., Hitchcock, C., & Menzies, P. (Eds.) (2012). *The Oxford Handbook of Causation (Oxford Handbooks in Philosophy)*. Oxford University Press. 51

Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, *23*(2), 287–298. 64

Bigaj, T. (2012). Causation without influence. *Erkenntnis*, *76*(1), 1–22. 52

Bigelow, J., Ellis, B., & Lierse, C. (1992). The world as one of a kind: Natural necessity and laws of nature. *The British Journal for the Philosophy of Science*, *43*(3), pp. 371–388. 44

Bird, A. (2007). *Nature's Metaphysics: Laws and Properties*. Oxford University Press. 44

Block, N. (2003). Do causal powers drain away? *Philosophy and Phenomenological Research*, *67*(1), 133–150. 58

Bohm, D. (1952a). A suggested interpretation of the quantum theory in terms of "hidden" variables. I. *Phys. Rev.*, *85*, 166–179. 47

Bohm, D. (1952b). A suggested interpretation of the quantum theory in terms of "hidden" variables. II. *Phys. Rev.*, *85*, 180–193. 47

Bohm, D. (1989). *Quantum Theory (Dover Books on Physics)*. Dover Publications. 54

Bohm, D. (1996). *The Special Theory of Relativity*. Routledge. 21, 202

Bohm, D., & Hiley, B. J. (1993). *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. Routledge. 31, 46, 47, 61, 92, 117

Bojowald, M. (2001). Loop quantum cosmology: IV. Discrete time evolution. *Class. Quantum Grav.*, *18*, 1071. 22

Bondi, H., & Gold, T. (1948). The steady-state theory of the expanding universe. *Monthly Notices of the Royal Astronomical Society*, *108*, 252–270. 64

Bourbaki, N. (1998). *General Topology: Chapters 1- 4 (Ettore Majorana International Science Series)*. Springer. 22

Bunzl, M. (1979). Causal overdetermination. *Journal of Philosophy*, *76*(3), 134–150. 68

Butterfield, J. (2005). Determinism and indeterminism. In E. Craig (Ed.) *The Shorter Routledge Encyclopedia of Philosophy*, (p. 195). Routledge. 20

Butterfield, J. (2011). Emergence, reduction and supervenience: A varied landscape. *Foundations of Physics*, *41*(6), 920–959. 151, 188

Callender, C., & Huggett, N. (Eds.) (2001). *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*. Cambridge University Press. 55, 203

Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford University Press. 44

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press. 151, 164

Chalmers, D. J. (2006). Two-dimensional semantics. In *Oxford Handbook of the Philosophy of Language*. Oxford University Press. 93

Clarke, R. (2005). *Libertarian Accounts of Free Will*. Oxford University Press, USA. 5

Collins, J., Hall, N., & Paul, L. (2004). *Causation and Counterfactuals (Representation and Mind Series)*. MIT Press. 51

Connes, A. (1994). *Noncommutative Geometry*. Academic Press. 175

Cresswell, M., & Hughes, G. (1996). *A New Introduction to Modal Logic*. Routledge. 96

Davidson, D. (2001). *Essays on Actions and Events (Philosophical Essays of Donald Davidson)*. Oxford University Press, USA. 151, 152, 157

de Broglie, L. (1925). Recherches sur la théorie des quanta (Researches on the quantum theory). *Annales de Physique (10) 3*. 46

Dennett, D. C. (1985). *Elbow Room: The Varieties of Free Will Worth Wanting*. Oxford University Press. 7

Dennett, D. C. (1997). *Brainstorms: Philosophical Essays on Mind and Psychology (Penguin science)*. Penguin Books Ltd. 5

Double, R. (1990). *The Non-Reality of Free Will*. Oxford University Press, USA. 5

Dowe, P. (1992). Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory. *Philosophy of Science*, *59*(2), pp. 195–216. 52

Dowe, P. (1995). Causality and conserved quantities: A reply to Salmon. *Philosophy of Science*, *62*(2), pp. 321–333. 52

Dowe, P. (2000). *Physical Causation (Cambridge Studies in Probability, Induction and Decision Theory)*. Cambridge University Press. 52, 63

Dowker, F. (2006). Causal sets as discrete spacetime. *Contemporary Physics*, *47*(1), 1–9. 51, 55, 87, 203

Dretske, F. I. (1977). Laws of nature. *Philosophy of Science*, *44*(2), 248–268. 44

Earman, J. (1986). *A Primer on Determinism (The Western Ontario Series in Philosophy of Science)*. Springer. 20

Eckart, C. (1926). Operator calculus and the solution of the equations of motion of quantum dynamics. *Physical Review*, *28*, 711–726. 46

Ehring, D. (1987). Causal relata. *Synthese*, *73*(2), 319–328. 54

Ellis, B. (2007). *Scientific Essentialism (Cambridge Studies in Philosophy)*. Cambridge University Press. 44, 77, 86, 116

Fair, D. (1979). Causation and the flow of energy. *Erkenntnis*, *14*(3), 219–250. 52, 63

Feynman, R. (2011). *Feynman Lectures on Physics*. Basic Books. 73, 110

Feynman, R. P., & Weinberg, S. (1999). *Elementary Particles and the Laws of Physics: The 1986 Dirac Memorial Lectures*. Cambridge University Press. 174

Fischer, J. M. (1995). *Metaphysics of Free Will: An Essay on Control (Aristotelian Society Monographs)*. Wiley-Blackwell. 6

Fischer, J. M. (2010). The Frankfurt Cases: The Moral of the Stories. *Philosophical Review*, *119*(3), 315–336. 6

Fischer, J. M., Kane, R., Pereboom, D., & Vargas, M. (2007). *Four Views on Free Will*. Wiley-Blackwell. 5

Forbes, G. (1986). *The Metaphysics of Modality (Claredon Library of Logic and Philosophy)*. Oxford University Press, USA. 96

Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, *66*(3), 829–39. 6

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, *68*(1), 5–20. 2

Frankfurt, H. G. (1982). The importance of what we care about. *Synthese*, *53*(2), 257–272. 2

Frankfurt, H. G. (1988). *The Importance of What We Care About: Philosophical Essays*. Cambridge University Press. 2, 7

Frankfurt, H. G. (1992). The faintest passion. *Proceedings and Addresses of the American Philosophical Association*, *66*(3), 5–16. 2

Garling, D. J. H. (2013). *A Course in Mathematical Analysis: Volume 1, Foundations and Elementary Real Analysis*. Cambridge University Press. 29

Griffiths, D. (2008). *Introduction to Elementary Particles*. Wiley VCH. 174

Halmos, P. (1974). *Measure Theory*. Springer. 29, 57, 63

Honderich, T. (2002). *How Free Are You?: The Determinism Problem*. Oxford University Press. 5

Horgan, T. E. (1993). From supervenience to superdupervenience: Meeting the demands of a material world. *Mind*, *102*(408), 555–86. 151

Hume, D. (2008). *An Enquiry concerning Human Understanding (Oxford World's Classics)*. Oxford University Press. 22

Isham, C. (1995). *Lectures on Quantum Theory: Mathematical and Structural Foundations*. Imperial College Press. 31, 117, 154

Jackson, F. (1986). What Mary Didn't Know. *Journal of Philosophy*, *83*(May), 291–5. 20, 30, 42

James, W. (2010). *The Dilemma of Determinism*. Kessinger Publishing. 5

Jammer, M. (1974). *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective*. John Wiley and Sons. 31, 46, 54, 117

Kane, R. (1985). *Free Will and Values*. State University of New York Press. 5

Kane, R. (Ed.) (2011). *The Oxford Handbook of Free Will (Oxford Handbooks)*. Oxford University Press, USA. 5

Kant (1964). *Immanuel Kant's Critique of Pure Reason*. Macmillan, London. 52, 125

Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays (Cambridge Studies in Philosophy)*. Cambridge University Press. 85, 151, 159, 162, 163, 166, 167

Kim, J. (2007). *Physicalism, or Something Near Enough (Princeton Monographs in Philosophy)*. Princeton University Press. 56, 58, 63, 189

Klemm, W. R. (2010). Simple experiments are not so simple. *Advances in Cognitive Psychology*, *6*, 47–65. 152

Kumar, M. (2009). *Quantum: Einstein, Bohr and the Great Debate About the Nature of Reality*. Icon Books Ltd. 46

Landsberg, P. T. (1984). *Enigma of Time*. Adam Hilger. 30

Laplace, P. (1902). *A Philosophical Essay on Probabilities*. John Wiley and Sons, New York. 18

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, *13*(4), 455–476. 94

Lewis, D. (1987). *Philosophical Papers: Volume II*. Oxford University Press, USA. 23, 67, 93

Lewis, D. (2001a). *Counterfactuals*. Blackwell Publishers. 44, 93, 95, 96, 97, 98

Lewis, D. (2001b). *On the Plurality of Worlds*. Wiley-Blackwell. 29

Libet, B. (1996). Commentary on free will in the light of neuropsychiatry. *Philosophy, Psychiatry, and Psychology*, *3*(2), 95–96. 149

Libet, B. (2005). *Mind Time: The Temporal Factor in Consciousness (Perspectives in Cognitive Neuroscience)*. Harvard University Press. 148, 149

Libet, B. (2006). Events and particulars. *Progress in Neurobioloy*, *78*, 322–326. 149, 152

Libet, B., Freeman, A., & Sutherland, K. (Eds.) (2000). *The Volitional Brain: Towards a Neuroscience of Freewill*. Imprint Academic. 149, 152

Loewer, B. (1996). Humean supervenience. *Philosophical Topics*, *24*(1), 101–127. 44

Lupher, T. (2009). A physical critique of physical causation. *Synthese*, *167*(1), pp. 67–80. 52

Markosian, N. (2000). What are physical objects? *Philosophy and Phenomenological Research*, *61*(2), 375–395. 174

McLaughlin, B., & Bennett, K. (2011). Supervenience. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Winter 2011 ed.
URL http://plato.stanford.edu/archives/win2011/entries/supervenience/ 151

McLaughlin, B. P. (1997). Supervenience, vagueness, and determination. *Philosophical Perspectives*, *11*, 209–30. 171

Mele, A. R. (2001). *Autonomous Agents: From Self-Control to Autonomy*. Oxford University Press, USA. 5

Misner, C. W., Thorne, K. S., & Wheeler, J. A. (1973). *Gravitation (Physics Series)*. W. H. Freeman. 54, 61

Nagel, E. (1961). *Structure of Science: Problems in the Logic of Scientific Explanation*. Routledge and Kegan Paul PLC. 188

Paul, L. A., & Hall, N. (2013). *Causation: A User's Guide*. Oxford University Press. 51

Paull, R. C., & Sider, T. (1992). In defense of global supervenience. *Philosophy and Phenomenological Research*, *52*(4), 833–53. 151

Pereboom, D. (2006). *Living without Free Will (Cambridge Studies in Philosophy)*. Cambridge University Press. 5

Pettit, P. (2001). *A Theory of Freedom: From the Psychology to the Politics of Agency*. Oxford University Press. 2

Poland, J. (1994). *Physicalism: The Philosophical Foundations*. Clarendon Press. 174

Popper, K. R. (1950). Indeterminism in quantum physics and in classical physics. Part I. *British Journal for the Philosophy of Science*, *1*(2), 117–133. 20

Price, H. (1997). *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time*. Oxford Paperbacks. 21, 30

Ramsey, F. (1990). (7) Law and causality. In *Philosophical Papers*, (pp. 140–163). Cambridge University Press. 44

Ramsey, F. P. (2013). *The Foundations of Mathematics and Other Logical Essays*. Martino Fine Books. 95

Redhead, M. (1989). *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics (Clarendon Paperbacks)*. Clarendon Press. 61, 203

Roskies, A. (2010). Freedom, neural mechanism and consciousness. In R. Baumeister, A. Mele, & K. Vohs (Eds.) *Free Will and Consciousness: How Might They Work?*, (pp. 153–171). Oxford University Press, USA. 19

Rovelli, C. (1998). Loop quantum gravity. *Living Reviews in Relativity*, *1*(1).
URL `http://www.livingreviews.org/lrr-1998-1` 175

Rovelli, C., & Smolin, L. (1995). Discreteness of area and volume in quantum gravity. *Nuclear Physics*, *B442*, 593–622. 22

Russell, P., & Deery, O. (Eds.) (2013). *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. Oxford University Press, USA. 5

Ryle, G. (1990). *The Concept of Mind*. Penguin Books. 124

Salmon, W. C. (1994). Causality without counterfactuals. *Philosophy of Science*, *61*(2), pp. 297–312. 52

Salmon, W. C. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science*, *64*(3), pp. 461–477. 52

Saunders, S., Barrett, J., Kent, A., & Wallace, D. (Eds.) (2012). *Many Worlds?: Everett, Quantum Theory, and Reality*. Oxford University Press. 92

Savellos, E. E., & Yalcin, U. D. (Eds.) (1995). *Supervenience: New Essays*. Cambridge University Press. 151

Schaffer, J. (1995). Contrastive causation. *The Philosophical Review.*, *114*(3), 297. 62, 63

Schaffer, J. (2000). Causation by disconnection. *Philosophy of Science*, *67*(2), pp. 285–300. 52

Schaffer, J. (2003). Overdetermining causes. *Philosophical Studies*, *114*(1-2), 23–45. 67

Schrödinger, E. (1978). *Collected papers on wave mechanics*. Chelsea Pub. Co. 46

Schutz, B. F. (1985). *A First Course in General Relativity*. Cambridge University Press. 21, 202

Seiberg, N. (2005). Emergent spacetime. *Rapporteur talk at the 23rd Solvay Conference in Physics, December, 2005*.
URL `arXiv:hep-th/0601234v1` 55, 203

Shannon, C. E. (1948a). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. 128

Shannon, C. E. (1948b). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 623 – 656. 128

Sider, T. (1999). Global supervenience and identity across times and worlds. *Philosophy and Phenomenological Research*, *49*(4), 913–937. 29

Sider, T. (2003). *Four-Dimensionalism: An Ontology of Persistence and Time (Mind Association Occasional Series)*. Clarendon Press. 29

Sklar (1992). *Space, Time and Spacetime*. University of California Press. 30

Smilansky, S. (2002). *Free Will and Illusion*. Clarendon Press. 5

Spivak, M. (2006). *Calculus*. Cambridge University Press. 29

Srednicki, M. (2007). *Quantum Field Theory*. Cambridge University Press. 65

Stalnaker, R. (1996). Varieties of supervenience. *Philosophical Perspectives*, *10*, 221–42. 151

Stalnaker, R. C. (1976). Possible worlds. *Noûs*, *10*(1), 65–75. 93

St.Augustine (2009). *Confessions (Oxford World's Classics)*. Oxford University Press, USA. 1

Stoljar, D. (2010). *Physicalism (New Problems of Philosophy)*. Routledge. 72

Strawson, G. (1984). *Freedom and Belief*. Oxford University Press. 5, 6, 124

Strawson, G. (1987). Realism and causation. *Philosophical Quarterly*, *37*(148), 253–277. 36, 38

Strawson, G. (1989). *The Secret Connexion: Causation, Realism and David Hume*. Clarendon Press. 23, 32

Strawson, P. F. (1962). Freedom and resentment. *Proceedings of the British Academy*, *48*, 1–25. 6

Suppe, F. (1977). *The Structure of Scientific Theories*. University of Illinois Press. 44

Swartz, N. (1985). *The Concept of Physical Law*. Cambridge University Press. 44

Tegmark, M. (2008). The mathematical universe. *Foundations of Physics*, *38*, 101–150. 44

Teller, P. (1997). *An Interpretive Introduction to Quantum Field Theory*. Princeton University Press. 174

Teller, P. R. (1984). The poor man's guide to supervenience and determination. *Southern Journal of Philosophy Supplement*, *22*(S1), 137–62. 151

Urbach, P. (1988). What is a Law of Nature? A Humean Answer. *British Journal for the Philosophy of Science*, *39*(2), 193–209. 93

van Fraassen, B. C. (1989). *Laws and Symmetry*. Oxford University Press. 44

van Inwagen, P. (1986). *An Essay on Free Will*. Oxford University Press. 4, 145

von Neumann, J. (1996). *Mathematical Foundations of Quantum Mechanics*. Princeton University Press. 46

Watson, G. (Ed.) (2003). *Free Will (Oxford Readings in Philosophy)*. Oxford University Press. 5, 7

Wegner, D. M. (2003). *The Illusion of Conscious Will (Bradford Books)*. MIT Press. 5

Weyl, H. (1999). *Space-time-matter (Dover Books on Physics)*. Dover Publications Inc. 54, 61

Yang, C. N. (1947). On quantized space-time. *Phys. Rev.*, *72*, 874–874. 22

Zwiebach, B. (2004). *A First Course in String Theory*. Cambridge University Press. 174