Emergent Wholes and the Porosity of Dynamic Objects

Submitted by Gabriel Arthur Thebolt to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Politics
in October 2013

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: …………………………………………………………..

Abstract

Claims in the metaphysics of strong emergence, featuring autonomous and possibly reflexive downward causal capacity, methodologically require, though ultimately ignore, units of analysis qua unified wholes. I argue that this avoidance of mereological and wider metaphysical debates denies the metaphysics of emergence clarity and cohesion and urgent application to conceptions of structure and agency. In this thesis, using a refined, non-linear, irreproducible, non-ontologically reductionist open-system physicalism and empiricism, I show that claims in the metaphysics of emergence hastily assume first the integration and subsequently the individuation of objects that become the subject of these strong claims. These assumptions, I believe, are actually the cause for the insurmountable gap between pure ontological reduction on the one side, and pure ontological and irreducible property emergence on the other. Furthermore, in using this new physicalism in the context of strong emergence, the traditional boundary between ontology and epistemology—going far beyond the standard weak-strong divide in the emergence discourse—can no longer be respected. As such, the nature of emergent properties is critical to assessing the nature of objects qua wholes with respect to the conditions for their integration and individuation. The major contribution to the metaphysics of emergence that this thesis provides is the realisation that, when we assume all physical objects are open and porous, all claims for persisting, emergent wholes are necessarily based on *physical assumptions of* integration and individuation. Synthetically I offer a method for understanding the individuation of 'quidditious' objects via properties when such a physicalist framework is employed.

Table of Contents

List of Figures & Illustrations

Introduction

The idea that there is something about the whole that is greater or more than its parts is not a new one, but even from a more contemporary perspective we often find that complex mixtures and constructs behave uniquely—perhaps even unpredictably—based on what we might confidently assume about their ingredients. “This such-and-such object looks, feels, tastes, sounds, and smells like something other than what it is made of.” Or, better yet: “This thing is doing things it should not be doing.” This is when emergence is relevant. It may be acceptable to leave it there, too, and just say that the odd behaviour is caused by some small detail that had been overlooked. But, within theoretical discussions, some have sought a more *metaphysical* understanding of such phenomena. Are we saying that the whole is doing something more than what its ingredients presumably allow for? If so, how is this possible? Or, are we saying that there is some effect of that whole that is only detectable when those exact ingredients are put together in that exact way? Or, more provocatively, are we saying that the whole *actually is* something more than its ingredients—that it stands apart from the catalogued sum of its parts?

0.1. Background

References to an object, person or whole’s supersummativity[1] (SS)—the possible condition of being more than what we know about its physical composition, or more than the aggregation of its parts—are found throughout the history of western thought often in the form of ‘the problem of the one and the many’.[2] It can point towards an ontological claim: an object simply *is* more than what its physical ingredients seem to

[1] This work is not rooted in the specifics of the problem of the one and the many. It is rooted in the problem of individuals. But, in order to arrive there in a coherent manner, and before we can begin to treat the problem, we need to see why the problem exists. Supersummativity is the most reasonable place to start.

[2] This phrase will be fully addressed in Chapter II. It concerns a classical question of how something can be both one and (made of) many things. The exact phrase is relatively recent, with particular usage by Peter Unger (1980), P. T. Geach (1980), David Lewis (1993), and Neil McKinnon (2002), with countless references in others.

amount to. It can also concern an epistemological claim: an object seems to be or do more than we what we would predict based on what we know about or expect from its physical ingredients. On the surface, these two types of questions are meant to be, and, traditionally are, separated. But, in the framing of some discussions of an object's supersummative features, it proves rather difficult to do so: is summation an ontological or epistemological issue? Could there be a sum, and could we know what it is for a given object? If we claim to know the sum of an object's physical composition, how can there be room for something other than the sum? On what grounds could we say that an object exists, of which we know its physical, catalogued sum, that also has properties disassociated from and irreducible to that sum? How can one be confident that they know the physical sum, when they have observational evidence suggesting there is more than what we had considered as the sum?

The classic dilemma present throughout metaphysics has always been that of consciousness or mind (in ontology) and personal identity (in epistemology): surely there must be something that makes me greater than a hulk of elements? I know that I, including my brain and its chemistry, am made of matter and force particles.[3] But, not only do I seem to maintain my own idea of myself from one day to the next, I also have the ability to move and affect change. Could there be something that is binding and organising me, giving me—or at least providing the means for determining—my own meaning and/or purpose, and possibly even some kind of central agency? So, we can ask whether a human individual or any object *simply is* more than its physical components, but we could also ask how or even whether such a question can even be answered. What is this sum, and how do we know it is such? What are we studying—what are we looking at—and do its properties result or emerge from (the knowledge of) that sum? I will contend that, with issues of SS, it is extremely difficult to gain any ground without allowing the threads of ontology and epistemology to become at least occasionally intertwined.

[3] Throughout this work, I will couple matter with force to connect my physicalist presuppositions to the standard model of particle physics, in which elementary particles are broken into two families of matter and four types of force carriers. Philosophically speaking, I think this is important for many reasons, but most especially because it allows physicalist analysis to recognise the inseparability of matter with force: one makes no sense without the other, even if only in the subatomic world. Because they are inseparable, fluctuations in both complicate our material definitions (and instantiations) of even common objects.

I will make all concerted efforts to respect these traditional boundaries, but because I am engaging with what is generally speaking the most recent attempt at tackling SS—the philosophy of emergence and all its different discourses—there may arise points in which these efforts evidently become futile. Emergence and SS justifiably are, or at least have inadvertently become, pan-philosophical issues, like existence and causation. So does that mean that the philosophy of emergence is really just a collection of various permutations of the metaphysics of emergence? If so, does this permit the crossing of lines in some cases? Let us unpack some of the emergence-SS basics first.

0.1.1. What is emergence?

'Supersummativity' is probably the best one-word synonym for emergence, which itself can be, conceivably, a phenomenon, a process, a cluster of properties, a position, a branch of metaphysics, a discourse, an informal or esoteric debate, or even a methodology. The word strikes a chord of *novelty*, possibly in or of form, essence, substance, mind, soul, configuration, causal powers, kind, property, and even ontological category. 'Novelty' could imply a lack of precedent, explanation, or prediction of any of these kinds of things, and the discourse, perhaps regrettably, touches on all of them. When I refer to novelty, generally speaking I simply mean *newness qua disassociation*: disassociation in property or kind from what seems to compose or cause it.

The philosophy of emergence, as it attempts to deal with the SS problem, is indeed primarily focused on novelty. This means, then, that, when taking into account all that the philosophy of emergence has to say on the SS issue, it is a discourse looking at novelty both in terms of existence and knowledge of such: summarily, it does, and, as I would argue, needs to, address both. So what are its origins? I'll start with what we might call explicit emergentism.

Self-professed emergentism—where the suffix indicates an identifiable movement, theory or persuasion—is a century-and-a-half-old school of thought on this SS issue (and, again, the nature of *summation* will be closely scrutinised in the pages to come).

In its simplest guise, the emergentists—those taking a 'specific' position on the SS issue—say that certain types of things that we study simply cannot be understood purely in terms of what we know them to be made of. There is something unique—novel, even—about the object we are studying, such that it has physical parts—it is made of physical things—but it also has an effect or feature that seems to be independent of them. In a similar sense, something new might be said to *come out of* what was already there (hence, in this sense, 'emerge', i.e. a process or a product of such), and it may account for things like self-awareness or self-organisation. It might be a new property (e.g. shape, colour, behaviour, configuration), or a whole new kind of thing entirely (i.e. a new kind of 'stuff', or a new category of existence). Some such modes are found in politics and sociology, as vast global structures and agents, markets and governments, cultures and languages, all seem to be and do special things in virtue of their scale and complexity. Some modes are also found in psychology and biology, as intentionality and cell mitosis, for example, seem to haunt pristine causal explanation and building-block composition.

As such a collection, emergentism *per se* first arrives in the late 19th century, largely in response to early (and successful) scientific mechanism—itself driven by the lingering Newtonian dream of causal and mathematical elegance—as a take on SS reformulated by John Stuart Mill in *System of Logic* (1843).

> All organised bodies are composed of parts, similar to those composing inorganic nature, and which have even themselves existed in an inorganic state; but the phenomena of life, which result from the juxtaposition of those parts in a certain manner, bear no analogy to any of the effects which would be produced by the action of the component substances considered as mere physical agents. To whatever degree we might imagine our knowledge of the properties of the several ingredients of a living body to be extended and perfected, it is certain that no mere summing up of the separate actions of those elements will ever amount to the action of the living body itself. (Book III, Ch.6, §1)

Despite numerous inaccuracies and incorrect assumptions in his chemical and physical analysis, Mill describes complex systems as having both homopathic and heteropathic laws and effects. The former concerns the mechanical, lower-level causal interactions: these are things we think we know well enough, like classical thermodynamics, where things seem to follow a unilateral, linear, apparently predictable course. The latter, however, concerns the association of these lower-level interactions as encapsulated by the system—which we can tentatively define here as a configured, multi-level unit of analysis[4]—and whether its (the system's) behaviour cannot be seen as an expression of, or be reduced to, the lower-level laws and effects. This incongruity between a system and its components (and their processes) proved to be the inspiration for several following him, including Alexander Bain (1870), George Lewes (1875, whose work included the coining of the term 'emergence'), Samuel Alexander (1920), C. Lloyd Morgan (1923), and C.D. Broad (1925)—together forming what Brian McLaughlin (1992) coined 'the British emergentists'.

This group, some of whom were also inspired by Bergson (notably *L'évolution Créatrice*, 1907), came to view existing scientific (and in particular, Darwinian) investigation as limited in its scope. Science had not only failed to explain the mind, but it had also not explained how certain objects behaved differently from what could have been expected of their parts. Mechanical science had given us many 'resultant' or homopathic effects: effects that are "the sum of what would have been the effects of each of their causes had those causes acted alone." (Lewes, 413) Yet, they argued, investigation had uncovered countless examples of 'emergent' or heteropathic effects: effects that are not the sum of what would have been the effects of each of its causes had they acted separately. In Morgan's *Emergent Evolution*, one example of 'an emergent' is chemical:

> When carbon having certain properties combines with sulphur having other properties there is formed, not a mere mixture but a new compound, some of the properties of which are quite different from those of either component. (1923, 3)

Though chemistry and physics surely contain emergents of this kind as well as many resultants, there was no reason to suspect much of the latter with regards to organic matter. And, this was more important for this group: the inability of physics and chemistry to describe and explain the creative and evolving nature of organisms and biological forms.

[4] The usage and definition of 'system' will be a key issue throughout this thesis, including whether or not 'system' can be used interchangeably with 'whole'.

The idea that there could be congruity—being constantly boosted by successes and consolidation in physics, chemistry and thermodynamics—has come to be associated with the methodology of what we now call reductionism. The more we investigate the natural world, according to this broad paradigm, the more we see how *A* is really just many little *B's* aligned in a particular way. Laws seemed to increasingly harmonise with long lists of constants,[5] 'confirmed' over many years of experimental testing.[6] Much to the annoyance of those trying to write textbooks, however, wherever science went, *incongruity* followed. One hand solidified, codified, and 'legislated' what previously had been hypothesised, as the other hand was digging up new dirt around the very same foundation. Emergentism—again, being a relatively coherent position on the SS issue—justifiably could not accept the flimsy promise of a complete reductive explanation as an *eventuality*.[7]

Today, much of modern science practice itself, with a little help from the history of cybernetics, non-linear dynamics, and what would eventually be considered as 'complexity' and its approaches, has now become adept at conserving some resources for the *unpredicted*, *unexplained*, and possibly *irreducible* features of certain kinds of objects qua 'systems'. Norbert Weiner (e.g. 1948, 1958), amongst his many achievements, showed how properties of a complex object ('system', in this case) can be fundamentally indeterminate due to the 'stochastic' distribution of its physical components and the perpetuation of system feedback.[8] Ilya Prigogine (e.g. 1961, 1977), also a widely celebrated contributor to a range of fields, provided some formalisation of what we can now call *open systems*: entities, including self-organising systems such as biological systems, in which matter and force are constantly being

---

[5] All constants do quite a bit of work. They represent generations and lifetimes of robust investigation, and are responsible for much more than just engineering achievements. But some constants change—or, are refined, as is preferred—whilst others are added—or 'discovered', as is preferred. What is more important, though: they are always limited to a finite number of decimal places.

[6] Even today, in the wake of the recent Higgs field 'confirmation', we still have much to learn about just how incomplete the standard model continues to be (something all physicists admit), with or without the elusive 'particle-field' alleged to give everything mass.

[7] There are some disciplines or sub-disciplines (e.g. particle physics) where some similar sentiments linger [e.g. Stephen Weinberg's philosophically-tuned *The First Three Minutes* (1979, 1993) and the more provocative *Dreams of a Final Theory* (1993)]. These are not only nomologically reductionist, but also ontologically and causally reductionist.

[8] In plainer terms, stochastic properties are statistically randomly (by a standard mathematical tests, and standard definitions of random, but still 'seemingly' random) distributed in an isolated system (force particles/fields only being exchanged with or interpenetrated by the environment). The feedback of the system concerns the transmission of information (and matter and force) in, through and out of the system in a gradual, perceptibly-stabilising process.

exchanged with the environment. The open system is the epitome of the dynamic system, with surprising and often novel features, and it is a concept that has continued to challenge the simplified and frequently reductionistic understanding of even common objects. Unmistakeably, the unpredictability and irreducibility of systems and system-level properties are, as we will see in Chapter I, two of the key attributes of the contemporary philosophy of emergence.

0.2. Contemporary discussions on the philosophy of emergence

Proposing and discussing the possibility of novelty, whether it be in substance, form, essence, accident, property, kind, causal capacity, or a number of other things—or, in other words, challenging the reductionist picture—has never really gone away in philosophy. Over the course of the last few decades, though, explicitly anti-reductive approaches have not only been a staple on the continental side of things, but have affected analytic philosophy as well. Jaegwon Kim went as far as to say that (apparently all) philosophy had "been under the reign of emergentism since the early 1970s." (1999, 5) As these sentiments fermented, exacerbated particularly by the slow crawl and apparent inadequacy of cognitive research, the critiques and ideas of complexity—vis-à-vis descriptions and explanations that could never be linear, static and reversible—gradually grew more diverse in application. The ontological and epistemological debates native to the metaphysics of emergence started to seep into the philosophies of sociology and social science,[9] typically revolving around claims about irreducible autonomous human agents and the spontaneous formation and development of the groups they compose. Simultaneously it became clear that the metaphysics of emergence needed also to converse with the philosophy of science and all its sub-divisions (e.g. philosophy of physics, biology, etc.), though presumably parallel ontological claims remain distinct from parallel epistemological ones. In a metaphysical system, however, there can be no objective division between these sub-disciplines if the analysis of the claims of emergence is in metaphysical language, yet relies on theories and units of analysis in any of these other philosophies.

[9] Notably, this includes Sawyer (2001) but also, in conjunction with structure-agency, Archer (1995) and Bhaskar (1998).

This thesis will not address the issue of human individuals qua agents in an explicit sense, as it sees the social forms of the claims for individuation, autonomy and SS to be rooted in what are ultimately metaphysical problems. This was, however, a preliminary motivation for embarking on this thesis: our structure-agency dilemmas in the humanities and social sciences[10] seemingly cannot be solved with discipline-specific methodological restrictions, and a philosopher of structure and agency finds herself in, or must at least engage with, the philosophy and eventually metaphysics of emergence.

0.2.1. Necessary terminology in the philosophy of emergence

Let us take another look at some of the language employed by the philosophy of emergence, as precision will become more and more critical as we move along—especially in the face of this ontological-epistemological challenge. Thus far, I have used ‘emergence’ largely interchangeably with ‘the philosophy of emergence’: the dedicated study of distinguishing in property or kind a higher level of analysis from a lower level, or the whole or system from its parts or constituents. The implication here is that this is a type of discourse, where the word suggests a host of positions and debates on the general question of whether SS is empirically or theoretically plausible. This of course differs from emergentism, which denotes a position, claim, or set of claims that tends to isolate and exalt the higher level of analysis. The British emergentists would all subscribe to emergentism, because they all conceive of, to take one example, irreducible features of at least one higher level. Hence, the word ‘emergence’, when understood as ‘the philosophy of emergence’, may simply make reference to or engage with the SS debate, whilst ‘emergentism’ must refer to one specific position or group of positions within that debate. A subset of the philosophy of emergence that will be the focus of this work, as well as the sub-discipline to which this thesis is contributing, is the metaphysics of emergence. This narrows the range of terminology, but expands the range of levels and units of analysis. I will address this further in the next section.

[10] As the thesis will unfold, I will at least hint at why it is likely that this is also the case in the natural sciences as well.

Today, the philosophy of emergence, a diverse, unsettled, and, frankly, chaotic, yet almost pedagogically-compulsory element of contemporary metaphysics,[11] pits *higher-level* analysis against *lower-level* analysis: both ends of the spectrum make descriptive and explanatory claims with which the other is often incompatible. This is the first set of parameters that we need to clarify here: in the contemporary philosophy of emergence, emergence and reduction are only relevant with an empirically-sound framework. This means we exclude any kind of philosophy that does not recognise the fundamental scientific principles of observation, description, explanation, and, for some, prediction.[12] And, this is largely how the philosophy of emergence operates, involving claims *about physical things* expressed in philosophical language—e.g. causal explanation, identity, persistence, etc. Because the sub-discipline of the philosophy of emergence relies on these scientifically-sound assumptions, we should therefore assume that claims about SS, whether they be emergentist or reductionist, of any persuasion, or anywhere in between, are to be assessed (as they predominantly have been) based on how they match with existing scientific observation, description, and explanation.[13]

Observation, description, and explanation (and, again, for some, prediction, or future estimation), when considered together in the philosophy of emergence, immediately remind us of what is ultimately the metaphysics of emergence and the inevitable mingling of ontology and epistemology: once we decide to evaluate claims about SS (or even summation itself), we are not only evaluating claims about what exists, but *how we know* it exists. Though every effort will be made to separate them, this is precisely why the philosophy of emergence is very much affected by these phases of

[11] All introductory courses and texts on contemporary metaphysics will include a section on emergence, reduction and supervenience. The metaphysics of emergence, the dominant and original flavour of the philosophy of emergence, is, especially when including these other two concept families, therefore its own sub-domain of metaphysics.
[12] This means that this thesis will not treat alternative ontologies such as substance dualism or pluralism, despite their role in the philosophy of mind and the emergentist threads therein—and the substantial portion of the philosophy of emergence that exclusively concerns the mind, volition, qualia, supervenience, intentionality, and a range of supersummative or potentially irreducible features.
[13] And, as we will see, they will be assessed also based on whether they accommodate the inevitable inconsistencies that come with human observation—inconsistencies that are themselves acknowledged by existing scientific investigation. This may not require a statement on the possibility for observation-independence or scientific realism. As we will see, a safer and less-contentious starting point is a refined physicalism and empiricism, to be treated in later sections.

investigation—and, as we will see later in this Introduction—the metaphysical currencies we require for optimal clarity. So, what are these phases?

Figure 0.2.1.1. 'Standard Investigation'

Investigation Phase	Definition
Observation	*Transmission of sensory data*
Description	*Coding of sensory data into properties* *{of: objects, events, other properties, conditions}*
Explanation	*Interpretation of (already-coded) properties* *{of: objects, events, other properties, conditions}*
Prediction (future estimation)	*Pattern ascription to (already-interpreted and already-coded) properties* *{of: objects, events, other properties, conditions}*

In Figure 0.2.1 above, the reader will notice the importance that an empirically-sound investigation places on the use of properties, as this will play a critical role in my engagement with the philosophy of emergence that follows in these pages. An important theme that I will tease out is the relationship between novelty in property and novelty in kind (of object)—and whether claims for novelty in causal powers can be seen to be derived from one, the other, or both. This is central to the reduction-emergence debate, and it is important to state at the outset the connection between properties and empirically-sound investigation. Additionally, whilst the precise definitions of the phases of investigation may vary within the philosophy of science, the general definitions given above do not contradict the scientific practice of how we might come to understand claims about SS in the philosophy of emergence. When referencing 'investigation' in this thesis, I am referring to these generalised phases—the defence of which I believe merits no further examination on the grounds that they are uncontroversial.[14]

[14] To be clear, in order to consider this as controversial, one would have to start by arguing that science practice does not involve observation, description, explanation and prediction (or future estimation)—which would not be a favourable position to have. These phases make no implicit comment on any form of scientific realism; they merely summarise how science practice operates. With regards to their description, I believe I have expressed them as concisely and clearly as possible. Ascribing a property 'red' (to an object, event, property or condition) requires an observation (red light being reflected and received), but also agreement on what 'red' means (hence, coding).

0.2.2. The philosophy of emergence in the context of structure and agency

The structure-agency debate, one of the most pressing issues in social science, dichotomises one singular unit of analysis, the agent, against a messy storm of influential forces and circumstances. Human behaviour—individual, institutional, or collective—is not immune to the effects of local physical, cultural, and ideational conditions.[15] Social investigation in a structure-agency context, and its vast range of methodological assumptions, must draw on what are ultimately metaphysical claims in order to translate the concepts of individuation, causal capacity and emergence for a given level of analysis. What are we studying, and what are we not studying? Is what we are studying something more than what it is made of? Is the behaviour of what we are studying inexplicable and/or unpredictable based on what we know about its composition? How far can our higher-level causal explanations go if we argue in favour of autonomy or ontological novelty? If confined to the social sciences, can we borrow or cherry-pick metaphysical arguments without extending the analysis beyond the individual human or group level? If cherry-picking, on what grounds might we make claims about autonomy or strong novelty in some form? These are only some of the issues at stake when the metaphysics of emergence is injected into the social realm.

But, structure and agency is a pandisciplinary problem, stultifying all kinds of investigation: is genetic expression the result of agency (the gene) interacting with structural forces (the physical-social environment)? If our analysis is confined to biology, we might very well argue that the answer is a timid ‘yes’, under certain conditions and using tentative and parsimonious definitions of ‘gene’, ‘environment’ and any other term we chose. Might we say that an organism ‘emerges’ from these objects, events, properties and conditions, that has autonomous causal capacity and irreducible system-level features? If we do, we certainly run the risk of dipping rather conveniently into the shallower waters of the metaphysics of emergence, which must, inextricably, flow into causation and composition.

[15] Many economic and political paradigms, in particular, attempt to minimise the role of ‘external’ influences in their descriptions, explanations, and most painfully, in their predictions. This is an example of a self-imposed disciplinary limit, and a major impetus for the broad scope of this work.

Structure and agency debates, regardless of discipline, at least *connect* to philosophical claims about individuation and causal capacity—which are claims that ultimately derive from metaphysical claims about those same things. Claims made within the philosophy of emergence, just as within the philosophy of structure and agency, are necessarily metaphysical ones: we are claiming to study *things*, that have properties, and that may appear to be individuated and possibly even autonomous. There is always a need to *study things*, and those things need to be distinguishable from the things we are not studying. If we are studying how a given financial market is affected by a given governmental policy, we must attempt to study *the market* and *the policy*. Despite our best efforts, our primary unit of analysis, the market, cannot be studied in a vacuum, nor can it be studied without limits on duration. We are forced to place limitations on *what we are studying*—and this is because a given social science discipline demands it. Should the policy prove to have prompted emergent properties in the market—e.g. unexpected macroeconomic behaviour—our explanations and analysis would have to state at the outset that they are limited to certain disciplinary boundaries. This is why the dilemmas in the philosophy of emergence—which as I hope to show encapsulates and supersedes the dilemmas in the philosophy of structure and agency—finds itself in a cul-de-sac when excluding metaphysical claims.

As with the philosophy of emergence, structure and agency questions are often pitched as necessarily self-confined to a given discipline. Limits on levels and units of analysis are self-imposed, such that conclusions can be drawn in the language of the discipline. One may assume that a human agent is individuated, because the discipline relies on using individual-level variables. Yet, the individuation and autonomy of a given agent is never assumed in the metaphysics of emergence and of structure-agency, which must grant some consideration of how we might know that (or 'why') *what we are studying* is actually *a thing to be studied*.

The primary point of intersection between the metaphysics of emergence and structure-agency is this very issue: what approach should we take in showing that an agent (or discrete, causally-efficacious object) can be studied in isolation from (or, as we will refine it as, 'is distinct from') its environment, such that we might begin to express, for example, their relations in causal terms? We know that an empirically-sound investigation must assume some degree of porosity despite needing *something* *to study*. Should, then, the philosophy of emergence attempt to break free of using isolated, static, causally-efficacious agents qua units of analysis? Or, should it grant some concessions to the uniqueness of societal-level forces, independent human actors, and macromolecular agents, despite, for example, the obstinate volatility at the macrophysical and microphysical levels? Philosophical claims for strong forms of emergence, with individuated higher-level autonomy, often featuring something called downward causation, must choose their battles wisely. Does it challenge the reductionist paradigm, but only up until a point, at the risk of being anthropocentric, and anti-metaphysical? Or does it make certain concessions to the ontological status of the individual person, whilst maintaining its causal powers? This is where the philosophy of emergence becomes the metaphysics of emergence: non-metaphysical claims cannot go any further, and the conclusions drawn in social science from the application of the philosophy of emergence—and, obviously, any hint towards any causal explanations—will be confined to that discipline. This is where the metaphysics of emergence takes over, existence, causation, identity, and persistence rear their heads, and where ontology and epistemology begin to blur together.

0.2.3. The metaphysics of emergence

As we will see in greater detail beginning in Chapter I, claiming novelty of any kind—objects, events, properties, conditions, possibly even new causal powers—is a very serious philosophical and, indeed, metaphysical claim. The stakes are very high—not just because someone might be, in certain cases, claiming a new kind of stuff, but also because the consequences, being metaphysical, trickle back into the broader philosophy of emergence, and eventually feed back into the philosophies of a range of specific disciplines. The philosophy of biology, to take one broad field, would be radically affected by changes in the basic approaches to observation, description, explanation, and, for some, prediction. All of these scientific trademarks are relevant to the metaphysics of emergence and all are affected by its claims.

The study of novelty, even in a 'weak' sense, must *at a minimum* attempt to treat complex phenomena first via description, as unpredictability, uniqueness, and blurred agency are all commonly observable. So common in our daily life are the strange and unexplained properties of ‘wholes’ or ‘systems’ under certain conditions that we rarely even notice them. The strange shape of a rock formation, the spontaneous building of ant-bridges, self-designing computer systems—things we are unable to explain perfectly—have to be *described* first. By observing and describing strange features, which seem to manifest themselves only after reaching a certain degree of complexity, one is identifying novel properties. Whilst the threads of a piece of purple cloth are individually purple, the molecules comprising the threads cannot be said to be individually purple (the chemistry simply provides for the purple light to be predominantly reflected—under common conditions). Hence, the piece’s purple hue is a descriptively novel property of a given unit of analysis. This is, rather simplistically, the aim of *weak emergence*: treating novelty in epistemic terms without any bold metaphysical claims. ‘Weak’ suggests frailty or timidity, and a weak claim is one where the stakes are lower.

Now, we could stop there, and say that every time we describe a whole (or at least, an object) we inevitably must include features not included in an exhaustive description of its parts. It could be simply the total mass of the particles or the rough shape of the rock or the general colour of the cloth. The aggregation or co-location of the constituents encourages us to individuate and ascribe system-level properties, or properties of the whole. The unusual shape of a coastal rock formation might seem unlikely to have come out of chunks of heavily-eroded minerals and ores: this would at least be a claim for weak emergence. The margin for error within such description is large, and the consequence for an error is minimal. Person A might consider it baffling and bizarre, Person B might consider it mildly intriguing, and Person C might consider it completely uninteresting because there are hundreds of other rock formations with equally random shapes and distortions—all manifested by a combination of forces over thousands of years. Nobody asked for an explanation of *how it came to have the strange shape that it has*. They just leave it there, and say, “The unique shape of that rock formation has emerged over time.” The shape—a description—emerges, and we might not care why. Stopping the investigation there, and confining it to observation and description, is weak emergence, as potentially characterised by downplaying the uniqueness as an indication of ignorance, or resulting from imperfect data collection, imperfect observation, environmental distortion, and a range of other possible causes.[16] This is metaphysical innocence, as Mark Bedau (1997) would say, and it is uncontroversial. It is not the metaphysics of emergence, if left there, and thus is not the deliberate focus of this thesis.

But, there is one huge problem with bracketing off weak emergence in the context of the thesis. In most cases, discussions in the philosophy of emergence also aim for the *explanation* of such phenomena, and any endeavour to explain emergent or system-level property X must be preceded by the individuation (and probably definition) of system S.[17] That is to say, the explanation of emergent wholes, systems, and properties stands on the shoulders of the description of these wholes and/or their properties. One cannot attempt to explain the existence of property X without having a description (and, therefore, at least weak individuation) of system S first. In other words, if we wanted to know why the cloth was purple, we would first need to agree that the whole's purpleness was interesting enough to merit further investigation—and the fact that there is incongruity between the properties of the cloth's molecules and the properties of the cloth's threads is significant. Furthermore, as we will see in the next chapter, taxonomies in the philosophy of emergence (from Kim, Robert Van Gulick, George Ellis and Terrence Deacon, Andrew Assad and Norman Packard, and John Searle) are *rolling* taxonomies: level two's features incorporate the features of level one, level three's features incorporate the features of level two, and so on. If description of novelty is level one, and indicative of weak or epistemic novelty, level two must proceed from the assumption that novelty has been observed and described, and incorporate the descriptively novel features of the system-whole.

Stronger claims about emergent phenomena would not dispute that there are some interesting features of the system that none of the parts individually have. The challenge begins when we wonder how these features have come to be. Is the system to be considered as a metaphysically-significant object qua unit of analysis, with

---

[16] Interestingly, the British emergentists (e.g. Morgan) contrasted an emergent (a noun) with a resultant (a noun), where the properties of the latter could be entirely expected based on knowledge of its composition.
[17] For much of this present of work, I will rely on variables or symbols—like O and P_1—not to associate myself with some kind of analytic fashion or style, or to prompt formal notation—but to denote the importance of *particular* objects as being things unique to the conditions of the thought experiment. Whilst it does improve a certain economy with words, my primary intention is to emphasise the irreproducibility of a given object, event, property, or condition, and the avoidance of universals and classes where appropriate.

system-level properties, a discernible configuration, or even having autonomy or downward causal powers? Is it a whole new kind of thing entirely, standing apart from the matter-and-force constituents? Is it unified, such that it is a unit of analysis from which nothing or no part is missing? Is it something to study that survives physical change well enough to merit a claim of persistence of some kind? If we want to go down this road, and begin asking a wide range of key metaphysical questions, we have entered the realm of (the metaphysics of) strong emergence—the primary terrain of this thesis. We are not simply looking to describe novelty, but we will also look at claims that go to great lengths to *explain* it, employing any and all metaphysical tools at our disposal.

If we wanted to explain why the cloth and/or its threads were purple, we would need to do more than *describe* the whole as being purple. Our question is: why is the whole purple, and the parts not (or why is the whole purple, *when* the parts are not)? The question is not *whether* the cloth is purple, but rather: upon agreeing that the cloth qua system or whole has a property ‘purple’, why, when we know that atoms and molecules could not have the property ‘purple’, is the cloth purple, when its microphysical constituents are not?[18] Is there a fundamental difference between the cloth (the system-whole) or even its threads, and its microphysical composition? Are there important distinctions to be made between the higher and the lower level? These are questions that result from attempts to *explain* novelty; and, their answers, if they are to be secured, must be found in the domain of strong emergence—or, the metaphysics of emergence.

[18] We would not only need a thorough account of the light-refracting properties of the atoms and molecules that comprised the cloth, but we would in principle also need to know the approximate conditions of and around both the particular cloth *C* and *C*’s observer. There would usually be no need to test the observer (assuming full faculties) or the environment (assuming standard conditions), both of which are subject to variation, but micro-variations in these and other aspects can never be overlooked in an exhaustive description and explanation. Though it would be strange and misleading to claim that we do not know why a particular purple cloth was purple in light of the achievements of modern physics—and an optical physicist would at least parsimoniously claim that we *do* know why *C* is purple—one can immediately see why explanation of novelty presents major challenges (with some cases of which possibly being outside the scope of any single discipline). I use ‘parsimoniously’ to indicate some of the assumptions embedded in modern physics, such as constants and so-called laws which can never be placed above or in front of observation. Physicist Paul Davies (2010) has written on the mutability of the laws and constants (as well as on emergence) in physics.

If we define novelty as any object or property that is both unpredicted (or unpredictable) and unexplained (or unexplainable),[19] then novelty can be described, and we could say that such a description is epistemological. This concerns how we know and how we have come to know system S qua emergent entity (e.g. a self-organising, autonomous whole). This also concerns how we know and how we have come to know property X qua emergent property (or emergent properties). Some of these emergent properties will seem to be causally efficacious (e.g. the ability X of system S to instantiate new properties in the world), whilst others are plainly descriptive (system S's shape X). All descriptions of S consist of linguistically-coded sensory data: an object we have identified as 'the cloth' carries a certain colour which we have called purple according to a particular linguistic tradition. We agree on what 'purple' is, and we agree on what object looks purple and what objects do not. The ascription of the property 'purple' makes sense because we have coded it as such.

In order to explain the existence of S (if S was a novel kind, or had novel causal powers, for example), or, in this case, S's system-level (i.e. disassociated, emergent) property X, I repeat: an object we have identified as 'the cloth' carries a certain colour, but we know that this means that the light that we will eventually call 'purple' is reflected by the molecules of the cloth back into my retinas, causing my brain to ascribe the property 'purple' to that which reflects purple light—the object we are calling 'the cloth'. Again, we agree on what 'purple' is, and we agree on what object looks purple and what objects do not (via repeated observations), but this is based on the reflection of the same kind of light to the retinas of multiple observers. In this way, we are explaining how the system-level property has been instantiated by lower-level properties (of objects, events, other properties, and conditions). This kind of explanation, and claims about such, is the primary focus of strong emergence as well as this thesis—notwithstanding its inclusion of description and its intersection of ontology and epistemology in the metaphysics of emergence.

[19] I add the alternative form parenthetically to indicate that there is an important contingent of emergentism that would deny the possibility of prediction or explanation in some cases—hence, 'able to be explained', as opposed to 'not-yet-explained' (or 'not explained as of current thinking', or 'not explained with existing data collection and/or interpretation').

0.2.4. Necessary currencies in the metaphysics of emergence

Now, as we are faced with pervasive novelty in our investigation into the philosophy of emergence, and since it can only be analysed in metaphysical language, we have yet another terminological way-point. The analysis which follows in this thesis will depend on our four phases of empirically-sound investigation, but also on carefully-chosen metaphysical currencies. These are devices that our metaphysical analysis of SS and novelty will be confined to, and they will be used in conjunction with the previously-mentioned phases of investigation. Further definition and justification for these currencies and their definitions will unfold as the thesis runs its course, hopefully minimising confusion in its argumentation.

Figure 0.2.4.1. Metaphysical Currencies

Currency & *Example*[20]	Definition & *Description*
Property	coded sensory data
M: redness *m: reflecting red light*	• *coding requires observation* • *corresponds to—and used to describe—coded objects, events, conditions and other properties (observation-dependent)*
Object	aggregates of matter and force particles/fields
M: apple *m: molecule*	• *individuated via properties* • *individuated by its ability to have a lesser degree of ontological and etiological dependency* than *other aggregates* on *other aggregates and the conditions they and it survive in (property-dependent)* • *existence implies suitability of conditions* • *requires instantiation for precise aggregation (property-dependent)*
Event	interaction of two or more objects
M: ripening *m: increase in sugars*	• *individuated via properties* • *event itself can generate event-level properties, which themselves can be property-dependent* • *interaction assumes instrumentally-discrete objects* • *'interaction' implies duration > 0* • *implies possibility of multiple levels of analysis*
Condition	topographical approximation of object behaviour and related events
M: 30°C *m: 9.81 m/s^2* ('g')	• *can imply difference in level of analysis relative to object(s)* • *often expressed in averages or ranges* • *topographical implies spatiotemporal, which implies duration > 0*

0.2.4.1. Physicalist assumptions

The reader will notice, from both the investigation phases and the metaphysical currencies, that the primary philosophical assumption in this thesis is not only one of sound methodological empiricism (in epistemology) but also one of physicalism (in ontology): I take the models and methods of modern science practice at face-value,

[20] M = macro-level (level of the system or whole), m = micro-level (level of the components or parts)

and assume that our starting point for any analysis is the assumption that the objects, events, properties, and conditions currently recognised and/or hypothesised by contemporary physics are the lowest possible level of analysis—and that *that* lowest possible level of analysis is conceivably the most appropriate for both description and explanation. In other words, we could (and probably should) explain, if not describe, in terms of matter and force particles, whenever we can.

On the surface, this is the opposite of what an emergentist methodology might be, is, or has been. It is, though, how all strong claims in the metaphysics of emergence actually make sense and prove significant: strong claims for ontological novelty or autonomous causal capacity, in their very nature, challenge the supremacy of microphysical explanation. The claims must stare into the face of physics and remain unsatisfied. And, this is the gauntlet through which strong claims in the metaphysics of emergence must run: with the stakes being high, in ontology, epistemology, causation, identity, and persistence, I believe I should start from the safest possible position, and that position, being roughly (or, in some cases, nominally[21]) the same position the bulk of the literature in the metaphysics of emergence starts in, should prove to be uncontroversial.

There is, however, a nagging problem with this presumably-safe starting position, whether or not it proves to be explanatorily sufficient in the context of strong claims: standard physicalism tends to conceal or downplay its ontological parsimony whenever convenient. This is to say, most physicalist starting positions in the metaphysics of emergence fail to admit that they are using some kind of primitive: ontological reduction requires an ontological primitive (an indivisible unit of analysis, or irreducible, ultimate bottom level of analysis), and yet this primitive must be identified by modern physics—and it is not. Physicalism in the context of the metaphysics of emergence, as we will see especially in Chapter I, recognises molecules, atoms, and even sub-atomic particles and the indeterminate implications of quantum mechanics, but only to challenge strong claims for any respective higher-level. In other words, physicalism, as it has been manipulated and conveniently

[21] I write 'nominally' here because not all positions described as physicalist actually account for all the lessons of contemporary physics.

reconfigured, simply dips a level down to challenge any claim for the failure of mereological supervenience (with *meros* as 'part', the claim of a higher level no longer being controlled by its lower levels).[22] But, it does this with an ace up its sleeve: this quasi-primitive, which could be anything from strings to quarks (depending on the scenario), within which everything and everything that modern physics could possibly hope to imagine or observe is contained, is simultaneously a lower-level (comparative), a lowest level (superlative), and an ultimate bottom unit (universal). In other words, if used in this way, physicalism becomes an obscurantist position and empirically untenable.[23] There are two major implications of this: 1) it gives physicalism an unnecessary and inappropriate ontologically-reductionist flavour, and 2) it leads physicalism to be accused, as I hope to explain in this thesis, of suffering from some of the same flaws as strongly emergent claims.

What needs to be said now, in this introduction, is that I intend to use physicalism—with its sound methodological empiricism and its guarded trust in physics and microphysical description and explanation—without convenience, without a bottom-level, and without an ontological primitive. I will attempt, over the course of this thesis, to recast a refined physicalism as a foundation without primitive and without absolute individuation. And, as we will see, claims made in the metaphysics of emergence will need to be revisited, as an uncompromising and occasionally inconvenient physicalism, which may lead to, among other things, mereological nihilism via uncharted ontological dependence, will eventually find wanting the exaltation of both the higher and the lower level.

[22] All of these terms will treated in due course. Mereology, the formal study of part-whole relations, and a focus of Chapter II, is a distinct, though rarely visited sub-field of ontology.

[23] Now, of course many have already noted this, and these exchanges are a key part of contemporary metaphysics and physicalist-realist ontological and mereological debates. Some good examples would include Lewis (1986) and Williams (2006). One of the resulting tensions from the idea that an ultimate bottom unit is more normative than empirical, is role of observation-dependence (O-D). This issue can be construed in a number of ways (and will be elucidated as we go along), namely: 1) radical: O-D where all objects of perception are entirely constructed; 2) standard: O-D where objects of perception can be real things, but our constructions or impressions of them are asymptotic and influenced by cognitive/technological filters. Though the latter is obviously the default position, of this author and the majority of modern science practice, they both can be oversimplified by the statement that 'nothing exists objectively'. The question "would the sun exist if we had always lived in caves?" is answered by "yes, but we might not have had a word for it."

0.3. Problems in the metaphysics of emergence

Despite its interdisciplinary and inter-scientific implications, when the examples are scarce, major authors in the metaphysics of emergence such as Kim, Searle, and van Gulick often rely on consciousness to push their models and analysis further. Many of these philosophers are, by trade, philosophers of mind, or, for the most part, metaphysicians of mind. The goal for them and many others is often to test various explanations of complex phenomena—the mind—using ontological, or strong emergence. How does the mind emerge from neural connections? Strong emergence says that the mind exists as something apart from those neural connections, a claim that must be accompanied by a thorough description and explanation. It is this question that for many major authors really drives emergence today.

This present work is not directly focusing on the emergence-of-mind problem, and this is for several reasons. The most important of these, and the only one worth mentioning here, is that, by exclusively looking at the mind, we risk excluding other complex phenomena. The intrigue of emergence is that it is a candidate—a quasi-foundational position—for an (metaphysical) explanation of the basic features of reality. I would pose the question, that if some principles contained within the metaphysics of emergence are feasible in one domain or on one level, why can or should they not be in or on another? Is the mind so special that it deserves its own metaphysical doctrine? Many of course would say yes, but such statements can often indicate a substance dualist predisposition, which will not be overtly considered here. Despite this deliberate avoidance of the chaotic yet often redundant emergence-of-mind area, I turn to a problem that encapsulates, eclipses, and, in my opinion, supersedes it.[24]

[24] In academic psychology, the so-called mereological fallacy (see Bennett and Hacker 2003, 2008) carries some favour. This is the notion that brains as units of analysis are insufficient for understanding how brains (or people) think. In other words, behavior strategy design is super-cognitive, and irreducible to synaptic patterns (which are themselves vertically irreducible). This is of course challenged on multiple fronts: a) the brain never was a whole in a metaphysically-significant sense, such that we could never treat it as such (which is precisely what I argue herein—see Chapter II); and b) but even though it is not a whole (in this sense) it can still be considered as the epicenter of all sensory data, and as such is where behavior strategy design is actually taking place. This thesis will argue—indirectly—that a human individual (or person) or her behaviour is not just more than a brain. It will argue that supersummativity was never quite satisfactory in this sense because there was never a possibility of sum. But, in order to arrive at that moment, we will need to let the claims in the metaphysics of emergence run their course—which as I will show needs to be mereological.

0.3.1. The omission of wholes *per se*

A slogan for emergent phenomena, philosophically as well as colloquially, might very well be that it indicates any case where 'a whole is (*or has features or properties that are*) greater than the sum of (*the features or properties of*) its parts.' This notion is trivially acknowledged at the outset of all of the most frequently-cited works on the topic, as if it were some compulsory mantra for any introductory framing of 'emergence'. And, for some authors in the philosophy of emergence, elucidating the term 'whole' from this mantra may very well be an unnecessary, possibly semantic digression—simply because there is no need to distinguish 'whole' from 'system' or higher-level entity. Emergent properties of the whole or system, whether they be weak (e.g. shape) or strong (e.g. causal powers), are being ascribed to a singular unit of analysis—the object that one is studying. The thought might be, 'Why worry about what we call it, if we have individuated it anyway, and we need to make sense of special cases (e.g. the brain qua mind)?'

The first problem that I am confronting is a mangled one: authors make assumptions about the nature of the objects they are studying, and these assumptions require them to (perhaps unwittingly) lean towards classical notions of 'whole'—which we might initially define here as a unified, higher-level, organised, individuated object. They then borrow these features of 'whole' and apply them to 'systems'. And, for the most part, when they reference systems, higher levels, and higher-level entities (and so on), they are really talking about 'wholes'. So, when Kim refers to a 'system', for example (1993a, 1993b, 1999a, 2006a, 2006b), he is really saying that his unit of analysis is a whole, with discernible parts and structure. This is, unfortunately, an enormous assumption to make, considering the mereological and pan-metaphysical stakes. But, that is not the biggest problem. The key issue with Kim as well as with the majority of the authors treated herein, is that there simply is no connection to the mereological *problematique*. There are evidently no resources spent on discerning whether the system does have discernible parts, unity, and structure. I am forced to ask: how can we examine claims about system-level properties and powers—which means properties of the whole—if we have not defended the assumption that the whole qua system could be individuated as system qua whole?

If the philosophy of emergence is indeed about wholes, or at least relates to them in some way, the discussion, if not analysing the nature of wholes *per se*, must include some assumptions about what wholes are. A whole has parts, or, in order for an object to be a whole, it must have, amongst other features, and depending on the mereology, parts. For a system to be a whole, it must have a physical composition and, presumably, some kind of configuration or organising force. Wholes and systems both have system-level properties, or properties of the higher-level. There is, by extension, a difference between the whole or system, and its parts or physical composition, but the former must encapsulate the latter. So, why the omission of the question of whether the thing we are studying is actually a unified thing to be studied? Why should we assume that the system encapsulates its composition? And, even if we did not go that far quite yet, why not at least acknowledge that the philosophy and metaphysics of emergence should have at least a working definition of system qua whole (or in other words, what 'whole' means)?

To study emergent phenomena, or claims about them in the philosophy and metaphysics of emergence, one needs to presume that the object to which emergent properties are potentially being ascribed (or the novel entity itself, in the case of ontological emergence) is actually *a thing to be studied*. An answer to the question 'is the brain qua mind a thing to be studied?' must precede the analysis of the question 'is the mental property X emergent from brain S?', even if the former garnered no reflection. One has, in this way, assumed that S is a whole, and that X is unexpected or unpredictable (and possibly even causally irreducible) based on what is contained within (the parts of) S.

Why attempt to examine claims about emergent properties or entities, when we have overlooked the very inspiration for distinguishing between higher and lower levels, between whole and part, or between microphysical (or microstructural) properties and system-level properties? The bulk of the literature avoids carving into the meat of the mereology of emergence, despite the latter being entirely determined by its assumptions. Objects qua units of analysis are examined, but objects qua wholes are not. This is our first problem to be addressed, and I believe the philosophy of emergence is a barren cupboard without its mereological dimension.

So what is so important about wholes, and why is it in the title of this thesis? If the metaphysics of emergence, in analysing strong claims, uses mereological terminology (which it does in this trivial sense), it (the metaphysics of emergence) could be, by extension, also a mereological debate (in addition to one of ontology, epistemology, causation, identity, and persistence). Mereology, being a fairly obscure branch of (often formal) ontology, immediately raises questions of wholes, parts, sets, and most importantly, ontological dependence or dependency relations. If the philosophy of emergence is broadly about the problem of the one and the many, we are, therefore, considering, first, objects which are individuated, and, second, objects which are unified (at least to some degree or in some way).

It so happens that, whilst the literature in the metaphysics of emergence only occasionally references ‘the whole’ *in the actual analysis*, it makes frequent reference to ‘the parts’. But, what is more important, there is no distinction made between ‘the whole’ and a range of what are probably synonyms for ‘the whole’, such as system, unit of analysis, higher-level entity, object, and so on. They are, when whole is even mentioned (again, it is never defined or treated as a concept), used interchangeably. And, as I will try to show in Chapter I, the authors are unwittingly justified in not making that distinction, as the metaphysics of emergence, in the analysis of strong claims, will indeed need to assume individuation, and, in some cases, unity and autonomy. It will become clear that they will need to have made assumptions about *what we are studying*, without any discussion of whether the thing we are studying is actually *a thing to be studied*. Pointing this out may be unnecessary, or capricious, or may result in a weak attempt to be grossly and hubristically panoptic on what should be a parsimonious issue. Furthermore, it may not be a question of whether X is a thing to be studied (individuation), but something less severe, such as whether X’s parts depend on objects not contained within the existing description of X (dependence). But, we should, I argue, intervene in the analysis of strong claims in the metaphysics of emergence considering the high stakes—in ontology, epistemology, causation, identity, and persistence. If a whole is *prima facie* that from which nothing or no part is lacking, then a whole is unified, and by extension, individuated. And, this is problematic, if only because there is no discussion on the nature of wholes in the existing literature in the metaphysics of emergence.

If, for example, claims in the metaphysics of emergence wanted to employ discrete units of analysis (individuated and presumably closed systems), which is actually the case with the majority of the strongly emergent claims, it would seem plainly obvious from the slogan above that the first two questions would therefore be:

- *What is a 'whole'?*
- *What are 'parts'?*

Strong emergentism entails the failure of mereological supervenience, such that there can be cases in which higher-level changes are not necessitated by lower-level changes. It is a failure of a doctrine that says that a whole—static or dynamic, discrete or continuous—cannot have features that have not been instantiated by its own parts. So, the claims of emergence must each at least generally rely on a fixed notion of what a whole and its constituents are. Yet amazingly, these two simple questions about emergence are largely ignored, even without a discrete or static presupposition *or statement of such*. Either the nature of the whole is in whatever sense taken for granted in some way, or a discussion on the nature of a whole is ignored because of an over-reliance on stock examples with embedded preconceptions of, for example, an object's persistence. What could possibly motivate a discussion of wholes and their parts more than a discussion of strong emergence?

In some cases, there is simply no mention of the whole. Authors such as Ernest Nagel, Kim, and Van Gulick question the viability of downward causation, for example, without really identifying what the agent might be or at least how it comes to exist. Many examples (often mind-related) are given, and new terms are introduced, but it is not clear at all why it would matter that we call the agent a whole—and therefore strongly emergent. Authors such as Terrence Deacon and George Ellis like to talk about 'levels' of a given unit of analysis. Yet, analysing the relations between levels must presuppose the coherence of the whole—and they fail to address it. In other cases, the *word* 'whole' is casually used, but only to diversify the range of synonyms for 'system', 'higher level', 'unit of analysis', 'object', and/or 'agent'. Maturana, Varela, Kauffman and others, though really not contributing to the metaphysics of emergence *per se*, at least make references to the unity of the whole, but they do not really demonstrate how the unit of analysis is actually unified or lacking no part or object. The assumptions these and other authors make about the integrity and material composition of the whole prevent what are essentially strongly emergent claims from holding as much weight as they might have otherwise. How can this be? How can we say that something is new, if we do not know what the ‘something’ is, or how ‘it’ comes to be?

In the metaphysics of emergence, authors rarely challenge the default matter-and-force physicalism.[25] None will deny the existence of the microphysical—some will even try to clumsily champion it to mask their proclivity towards ontological reductionism. Some will try to show how existing microphysical models fall short of explaining, for example, the apparent autonomy of a given higher level or unit of analysis. But, here is the problem that seems to blind them from their metaphysical omissions: claims in the emergence literature are constantly made *about physical things*, but they are claims in which the authors chose *which* levels of physical things to consider. Some things have been left out, whether it be by convenience, alleged relevance, or, conceivably, ignorance. If, for example, a philosopher of biology wants to talk about the developmental plasticity of given class of eukaryotic cells, we could expect her to confine her descriptions and explanations to biological structures and agents—a discussion on biology involves levels and units of analysis appropriate to the discipline. At the same time, though, if *anyone* wanted to make claims about ontological novelty and autonomous causal capacity, *there would be no boundaries to the analysis*. And, if the analysis is unsatisfactory, it is probably because certain assumptions were made, and certain structures and agents were deemed off limits or beyond the scope of the work. Such claims, regardless of their literary economy, are still unsatisfactory.

The philosophy of emergence (and, what is more important, the metaphysics of emergence), as we will see, presupposes the physical, for if there were no physicalist base to start from, novelty would not be impressive. We start with the physical, and

[25] This will be a recurring theme of the work. Physicalism has many connotations, chief among them is reduction. Real, contemporary physicalism says nothing about actual ontological primitives, because physics does not require them and because physics does not have agreement on whatever they might be. Therefore physicalism cannot be ontologically reductionist, unless a given proponent has misunderstood contemporary physics. The appropriate term, I would think, is property-ontological reductionist. Chapter I will open up the discussion on this issue.

see how these claims fare. Again: how can there be claims for novelty, if we whimsically ignore not only the mysterious nature of the quantum world, but, as we will see in some cases, basic atomic physics, chemistry, thermodynamics, and chemical and cellular biology? Microphysical structures and agents are never irrelevant, and because they are not, they could never be inconvenient: they are there, and we must incorporate them. The lack of precision about what we are studying, when coupled with bold claims of strong emergence, makes the philosophy and metaphysics of emergence even more clouded and confusing than it already is. If we are to test objects for possible emergent features, we need to be clear on what it is we are studying. Do we have a whole or not? Is there unity or not? Are there discernible parts or not? Answering these questions in a metaphysically-significant sense precludes the application of emergentist ideas across the social sciences and is not afforded the luxury of excluding inconvenient factors, forces, particles, fields, objects, events, properties, and conditions. Emergent properties and entities must be seen through an uncompromising physicalist lens (or, treated by a refined physicalism[26] and empiricism) first—especially when the emergence literature shows no evidence of having considered that from which novelty is manifested—the 'whole'.

If emergence were to provide, for example, a rigorously organised set of legitimate options on the nature of the whole, working with a physicalist base, many elements of the philosophy of emergence might prove to be quite an attractive means of describing and possibly explaining human reality in metaphysical terms. That means that the motivation for describing and explaining (or at least analysing) phenomena through an emergentist lens is well-intentioned and far from being presented here as unhelpful—because it reflects some awareness of the evidence for the dynamic and at least non-ontologically reductionist nature of physical reality. The other candidates, the challengers to emergentism, do not have a strong track record, according to a new

[26] There will be many references to this term throughout this present work. I use it in order to distinguish itself from plain physicalism (often implying, though not requiring, an ontological primitive). This refined version signifies a reverence for physics, of course, but it also signifies the inherent non-linearity, irreproducibility, and metaphysical uncertainty that modern physics has established. So, there is no bottom unit and no ultimate causal explanation or model. There is, however, the assumption that all things are made of physical things, and by extension, the emphasis on upward-looking compositional determinacy and causal explanation. Considering the microphysical volatility that we will see in the chapters to come, a whole's persistence is fundamentally tenuous. As such, a refined physicalism implies upward explanation with compositional dynamism and local condition-dependence. Again, this is a key part of this present work, and will be referenced throughout.

physicalist (or a new naturalistically metaphysical) standard. Pure ontological reduction was seriously challenged by the Copenhagen interpretation of quantum mechanics[27] and the subsequent, as-yet unsuccessful attempts at a theory of everything (or unified model). Its twin—atomistic physicalism—is dependent on a standard model that cleverly masks its confusion over gravity and mass, and for some, also depends on an unobserved ontological primitive (e.g. string).[28] Making sense of wholes, as a necessary preamble to analysing claims about strong emergence, must go beyond the legacy of conveniently-static, closed, and isolated units of analysis that has confounded most discussions of supersummativity.

0.3.2. The omission of wholes as persisting

Our first problem concerned the absence of a treatment of 'unit of analysis qua whole', so this indicates that the metaphysics of emergence fails to operate with a working definition of whole (or a choice from within the taxonomies of such, to be expounded in Chapter II). Without a system in place to answer the basic questions of 'whole', including what the criteria for 'whole' might be, how it might be formed, how it might come to be known, and so on, it will prove difficult to draw any conclusions on much of the analysis of strong claims. Furthermore, by not knowing what a whole is or what one might be, we should also presumably be unable to justify the individuation of such—and the system-level (including causally-efficacious) properties that might be ascribed to one.

In addition to, and as an outgrowth of, the omission of wholes *per se*, scrutinising claims of strong novelty is also contingent upon the nature of the whole's identity and

[27] QM essentially presents us with the anthropic principle: observations, which lead to hypotheses and theories, are dependent on the observer and her technology. QM does not emphatically discredit causal reduction at all. In some sense it is its greatest champion yet (at minimum via probability matrices). But QM does demonstrate a fundamental uncertainty of the ontological status of the things included in the standard model. Many different particles are doing many different things, all whilst having properties (evidently) dependent on observation. Subsequent attempts to construct a theory of everything such as string or M-theory—which would provide an ultimate bottomed-out unit—have yet to be accompanied by observational evidence.

[28] Some philosophers, noting the infinite regression problem, have postulated something called 'gunk' (an infinitely-divisible, atomless object having only independent parts) in competition with mereological nihilism (the doctrine that certain objects never compose any further part). I will give this some treatment in Chapter IV and in the Conclusion.

persistence. All of these authors above (and many more) have evidently been compelled to overlook, amongst other things, certain kinds of microphysical change and its implications for individuation—and this omission, once again, is not permitted by a refined physicalism (see note 23). Even from the time of the early emergentists, authors in the metaphysics of emergence, as a result of failing to define 'whole' and 'parts', have also not addressed these fundamental mereological questions:

- *How does a whole come into existence?*
- *How can a whole be identified?*
- *How can a whole change?*

Asking what a whole and its parts are leads one to ask how they are to be identified. In order to study a whole—and examine strongly emergent claims for them—we to need to know how it has come to be (which will affect our decision to consider a system S as a whole). Does the whole create and sustain itself in a metaphysically significant sense, or in a sense that merits strong claims—possibly even of autonomous causal powers? This is initially a mereological question, as the answer will be determined by a specific mereology (and, more broadly, a metaphysical system) with rules of composition and dependence. How is the whole made, and how do the parts relate to the whole? When we test a given object S for whole-ness, and its compositional origins, we depart from the mereological and shift slightly into the empirical. If (based on a given mereological position) the whole or higher-level determines what parts or physical composition the system can have to make it such, do we have observational evidence for confirming this? Has a given object S made and sustained itself, or at least become something new (or interesting) in virtue of its higher-level? If we are satisfied with this clarification, regardless of the answer, we should be able to move on the next question—though this formational question should be asked first (after deciding what a whole could be).

Next, we need to know if we are justified in *individuating* such a whole first. What are we looking at, and what are we *not* looking at? Can an object S be studied in isolation from its environment? If system S is composed of small, fluctuating things, some of which come and go, can we say that we know what S is? What are we studying, if we admittedly only have some of the properties of S, which themselves are subject to coding and interpretation? In the social sciences, one is often expected to individuate the human actor. In the metaphysics of emergence, we are not offered the luxury of setting aside certain objects—especially biological systems and the mind. Are we justified in claiming, for example, endogenous, self-organising, or autopoietic novelty?[29] The discourse must consider whether strong claims are even on the table if the individuation of the object qua system-whole is uncertain.

The identity problem is the preamble to the persistence problem. All units of analysis, in a physicalistic picture, normally confined to one discipline or another, find no respite from constant property alteration—with or without the assumption that the properties we *have* ascribed are all of the properties that *could have possibly been* ascribed. What we are observing at this moment, cannot possibly be precisely what we are observing at another, if only referring to microstructural or microphysical properties. With change being inevitable, at least at the microphysical level, a whole's persistence becomes a significant obstacle to analysing strong claims.

If we claim to know what object we are studying, and that that object can be studied in isolation from its environment, how does the object survive change? 'Survive' and 'change' are loaded words. Survival of an object implies, first, duration: something that survives exists at both t_1 and t_2. Survival also implies shared properties: an object is the same or similar at both t_1 and t_2. Change implies, in addition to duration, the confounding combination of similarity and difference: change requires that an object be the same *and* different between observations, as we will see in Chapter IV. So, is the object the same, such that we are justified in claiming that an object is strongly emergent at both t_1 and t_2? The only way we can answer these questions is to dip into the persistence literature, in which there are two general streams that in some ways neatly pair, I think, with the mereological options outlined in Chapter II.

If we claim that an object S is a whole, and that that whole can be individuated, can we say that that individuated whole is the same whole at multiple observations? One persistence camp says yes via endurance: wholly persisting, without change to

[29] Self-organisation and autopoiesis will be treated in Chapter I §2.7. Many would consider this an odd question, akin to asking whether a human individual is distinct from the floor on which it stands. I would encourage such a reader to consider carefully the case studies in Chapter III.

('important' or *essential*) parts. This means that there is something about the whole that exempts it from what might be trivial microphysical forces in light of configuration, organisation, unity, autonomy, consolidation, durability, and even form, substance, and a number of other possibilities. The other camp says yes via time-slices: an object S is the *a posteriori* aggregate of all the S-at-t_x's. In other words, very briefly, if we said that S existed both at t_1 and t_2, even with physical changes, *S per se* is the amalgamation of *S-at-t_1* and *S-at-t_2*. It would be like saying that I am now, what I am now, plus what I was earlier. On the surface, this may seem appealing, but we shall see whether this is appropriate for an uncompromising physicalism, and, most importantly, whether this is appropriate for strong emergence.

0.3.3. Summary of problems and their urgency

Chapter I is dedicated to excluding the possibility that strong claims in the metaphysics of emergence do not need a whole, and, in the process, presents a wider understanding of emergentism and reductionism. Since claims about strong emergence seem to require a working definition of whole-ness, which is not provided, I consider the presentation of 'whole' to be essential for the metaphysics of emergence. Options should be provided for what a whole can be, and only after that might strong claims be analysed. As it turns out, there are a few more steps involved in doing so.

If we have an idea of what a whole is (Chapter II), or what the criteria are for an object to become whole, and therefore eligible for emergent ascription (as we shall see in Chapter I), we can decide whether it is possible that an object S could indeed be a whole. Part of *that* decision is influenced by any evidence we might have to select that object as being our unit of analysis—whether it be qua whole, system, or just a an unorganised heap of matter and force particles. Depending on the option, and the mereology one has reason for choosing through a physicalist lens, we can consider how and why such a whole has come to exist, and whether it can be individuated. If individuated, does the object endure, perdure, or persist in some way at all? By being uncompromising with our physicalism, where no bottom level is assumed, and no level is off limits, how will these options fare? Will objects survive, so that we may be justified in defending stronger claims?

It might be the case that, despite the incessant microstructural property alteration, a novel causal capacity of a given higher level, for example—even if not meriting a new category of existence (e.g. substance dualism or ontological pluralism)—actually gives us reason to view the fluctuation as resulting from the inevitable porosity all physical structures must have. Perhaps all physical aggregates are porous—meaning compositionally subject to environmental influence—and yet can do special things at the level of the whole despite this influence. If this is the case, it could mean that the whole, in virtue of some kind of quest for equilibrium or even-distribution of energies, was re-inventing itself whilst maintaining many of the same micro- and macro-structural properties. Many of these questions will simply not be *answered* in this thesis, but I hope to at least *present them as problems*.

0.4. The thesis

From the outset, I expect the following propositions to be true:

- **(0)** *Candidate-objects being emergent or having emergent properties must be wholes.*

- **(1)** *If we do not know what 'whole' is, we cannot say that object S is a whole.*
    - *1a) If we cannot say that S is a whole, we cannot say that S or its property X is emergent.*

- **(2)** *If we cannot say that S is a whole, we cannot say that S persists.*
    - *2a) If we cannot say that S persists, we cannot say that S can be a whole.*
    - *2b) If we cannot say that S persists, we cannot say that emergent property X can be ascribed to S.*

Claim (0) is the focus of Chapter I, and I have included it to cover my bases and to justify such a lengthy discussion of emergent wholes. Claim (1) is presented and initially defended at the end of the first chapter, and (1) and (1a) are twinned with (0) in Chapters I and II—though ultimately are emphasised throughout the work. Claim (2) is initially presented in Chapter III, and (2a) is the major focus of Chapter IV. The end of Chapter IV presents Claim (2b) in what I will continue to refer to as a refined physicalism and empiricism. Though admittedly taking these claims at least initially to be uncontroversial, I recognise the necessity of demonstrating why they are critical to the metaphysics of emergence.

0.4.1. Contribution

The philosophy of emergence casts a spotlight on the new, the strange, the previously-unobserved, the unpredicted, the non-reproduced, the incompatible, the unreduced, the unexplained and the creative. These things are there for us to see, and there is no denial of their existence. It is at this point where emergence diverges into weak and strong forms: weak emergence leaves these things as they are, as it does not make holistic or even metaphysical claims, and is content to preserve them as descriptively novel. The *strong emergentist* goes beyond description, and converts them, as corresponding to the above adjectives, into the intrinsically unique, the unpredictable, the irreproducible, the fundamentally incompatible, the irreducible, the unexplainable, and/or the autonomous. The *strong reductionist* converts them into the old-in-a-different-form, the temporarily-unique, the predictable-in-hindsight, the eventually-reproducible, the compatible-with-caveat,[30] the reducible, the explained-away, and the configurational.

My work here dissects these two strong variants as they have been presented hoping to show why they are both well-intentioned and occasionally promising, but ultimately too inconsistent and fragmented to do any real explanation. The dichotomous literature on strong emergence and reduction, relying on discrete and often static and enduring units of analysis, overlooks the nature of wholes and their persistence. I simply show why the philosophy of emergence must discuss wholes, and when it does, how poorly the claims about strong reduction and emergence fare. Gradually, over the course of

[30] This is to say, they would claim it is compatible with a standard model, provided the standard model is correct or eventually completed.

the four chapters herein, I try to synthesise some aspects of both to make the metaphysics of emergence actually compatible with 'whole' and worthy to take on the problem of supersummativity—with potentially broader, philosophical consequences.

0.4.2. Methodology

It is my aim to work exclusively within the area of the metaphysics of emergence, and only after this might the philosophy of emergence be better applied and incorporated into a wider range of philosophies in the natural and social sciences. Because of this, the thesis restricts itself to four sub-domains of metaphysics (as they connect to claims about emergent properties and entities): ontology, epistemology, identity, and persistence (causation in the philosophy of emergence is of course also considered as a sub-text of all four). I treat them in this order, where the first two coincide with the first two chapters, and the latter two for Chapters III and IV, respectively.

The starting point is the emergence literature: if major authors like Kim, van Gulick, and Searle (and others) want to examine claims about emergent properties, powers or entities, they will need to declare what a whole or system is. Because they do not, I provide close analysis of their writings to show that, despite using a range of terms, their units of analysis are strikingly dependent on whole-like assumptions. This allows me to connect the metaphysics of emergence with mereology and its history of concepts. I follow the trail of conceptualising 'whole' and its ontological-identity questions via a range of key contributors to the problem of the one and the many. By presenting this brief genealogy, I am able to frame the struggle that the philosophy of emergence has with understanding 'whole', such that we can see how and why the metaphysics of emergence has been until now unequipped to deal with the very ontological-epistemological intersection it has created.

In other words, I will attempt to show that the status of a whole's existence (and, by extension, its individuation, and any emergent powers and properties) is entirely dependent on how a whole is determined (both in terms of causal powers and compositional determinacy). In a return to close readings in the metaphysics of emergence, we can ask, simplistically, does a whole make the parts, or the parts make the whole? This means that in order to examine claims about emergent powers and properties, we need to agree on how the wholes have come to be and how they have come to be known. As this unfolds, from providing a range of definitions or features of wholes, to showing how something might be *considered* a whole, it becomes increasingly difficult to supress the nature of the whole's persistence, and whether a whole at one time can indeed be the same whole at another. After asking the question of what a whole is (II), and how it can be determined (III), we need to know how our concept of whole confronts the possibility of change (IV). I engage with the metaphysics of persistence—and its two primary branches of endurance and perdurance—and see how our definitions (II) and composition-causal pathways (III) might survive over time.

From beginning to end, I stand on an empirically-sound foundation, clearly distinguishing between the four phases of investigation and the process through which properties are ascribed. By doing this, I remove many of the standard restrictions on levels of analysis despite maintaining the same metaphysical currency-language throughout. The establishment of these currencies, and their consistent use, should, hopefully, make the transition between larger-scale objects (e.g. mountains or markets) and smaller-scale objects (e.g. bosons and bacteria) less challenging.

A physicalist lens is used, and no exemptions for microphysical uncertainty and fluctuations have been included. This, as it turns out, I believe, should not be controversial, as the bulk of the literature in the metaphysics of emergence is already starting with the same (albeit, fractured) foundation. What may be considered as a pre-existing penchant for incessant metaphysical vagueness (in both the technical and general sense), should not be challenged, I would argue, on the grounds that the stakes are so high with these metaphysical questions.

0.4.3. Chapter breakdown

First I go to some length to show why emergence (along with the inverted reductionist counter-arguments) is really about the whole, and why references to system, higher-level, object, unity and totality all converge on the notion of a singular unit of analysis.

Chapter I starts with the primary features and forms of reduction, with which the contemporary philosophy of emergence is trying to come to terms. Here we might try to find some merit in *property-ontological reduction*, though not pure ontological reduction. But, from the majority of these assumptions we see a hint of the same problem to come in the next section, which breaks down the fundamental doctrines of emergence itself. Looking first at the ontological dilemma of the metaphysics of emergence and its weak-strong divide, then moving on to some of the more celebrated current work on emergence, Chapter I's purpose is to confirm the need for wholes in any discussion of these strong claims.

Working on the assumption that wholes qua units of analysis are necessary for any such scrutiny, Chapter II connects emergent wholes to classical mereology and theories of unities. From Classical, Scholastic, and modern treatments of 'wholes' and their parts, we can begin to see how the term 'whole' came to be indicative of supersummativity: something, as in, for example, *form* or *essence*, must be there to bind, organise, and possibly even 'teleologise' the physical components—but certainly in order to associate, collectivise and, most importantly, individuate. In general, the history of mereology yields, by my account as well as others, four themes that the selected contributors can be identified by: (1) the status of the whole's unity; (2) the ontological categorisation of both parts and whole; (3) dependency relation between parts and whole; and, (4) relevance of order and structure of parts to whole. Depending on where the authors stand relative to these issues, the themes seem to point to a key polarisation of macro-micro compositional determinacy and causal powers. Additionally, a significant development in mereology will be shown to be critical for emergence studies: the *integral whole*, originally a medieval addendum to the Aristotelian taxonomy of essential and aggregate wholes, is meant to demonstrate how some units of analysis have compositional fluctuations over time whilst arguably remaining wholes.

Sometimes the whole makes the parts, and sometimes the parts make the whole, depending on the mereology one is following. Chapter III takes the micro-macro dilemma relative to both composition and causation and tries to see how the claims of strong emergence fare. If we have some kind of supersummativity, what are the implications for, let us say, downward causation, in a physicalist picture? Do we have an understanding of summation that would allow us to not only individuate an object, but to ascribe novel properties to it? Here I provide a much more precise treatment of downward and higher-level autonomous causal capacity, in light of what we now can postulate about the whole. Three extensive cases are presented that demonstrate the fundamental misapprehensions of the nature of *the units of analysis qua wholes*—and not vice-versa. One key realisation here is the idea that—using any mereology or directionality— the individuation of (at least lower-level) units of analysis qua wholes is dependent on these same issues of compositional determinacy and causal capacity. In order to claim strongly emergent features, we would need to identify and isolate the object with the knowledge that doing so would require the determination (and construction) of the whole. In a metaphysically naturalist picture, this is a daunting task due to both the fluctuations in local conditions and, by extension, the volatility of its composition.

The examples from Chapter III show that constructing a whole—which, depending on the mereology being used, flows macro to micro or micro to macro—must apparently make assumptions about the whole's persistence given the uncertainty of the local microphysical conditions. Yet, this is hardly something to overlook. Chapter IV begins by showing that individuating the unit of analysis qua whole (or vice-versa) must always precede determining if and how it has changed over time. As the previous chapters had shown, this is already a time-saving, instrumental technique that ultimately only stultifies the metaphysical analysis. The nature of the persisting object's change must depend on that object being both the same as well as different between two observations, meaning that microphysical properties inevitably would prove to supersede whatever system-level properties we thought we had ascribed. With the various doctrines of temporal parts seemingly appealing to the non-spatiotemporal, and the viewing of time as a discrete fourth dimension, four-dimensionalism (worm perdurance in particular) provides only some help in showing the compositional dynamism of wholes qua units of analysis. One major question is whether a *quiddity* (what-ness) of a unit of analysis can be generated from anything other than a supersummative mereology. Two key concepts offered here are the *entia per alio* and the *entia successiva*, notably in a reformulation of the classic *Ship of Theseus* question. In general, though, Chapter IV concerns itself with how a whole qua unit of analysis can change—considering its inevitable and constant property alteration—and whether autonomous and irreducible features can justifiably be claimed in light of such volatility.

Within the title and throughout the work, the phrase “emergent wholes” is used. This is because the objects often employed in the discourse—the examples used to show whether or not something has emergent features—are discrete and static or exclusive of certain levels of analysis. Ascribing emergent properties, just as with identifying wholes, requires unity *of some kind*, despite the seriousness of claims for metaphysical integrity. This tension—between needing things to study, and knowing whether the thing we wish to study is actually a thing to be studied—is critical throughout. Because from the beginning there is a physicalistic assumption, this fundamental dilemma, I believe, encourages us to blur those ontological and epistemological lines, and question the identity and persistence of the very objects we wish to consider as emergent.

The goal is to start from where emergence starts (physicalism), and see how the strong claims fare if they had been exposed to a discussion of ‘whole’. At the very least, I aim to show how the ideas in the philosophy of emergence—and all the novelty and supersummativity that they wish to point to—might be better understood and applied after such an exposure. And, in the final synthesis, I hope to show what an emergent whole might be, how it might come to exist, and how it might survive change.

50

Chapter I | Does emergence need a whole?

Perhaps the root cause for concern for me in the metaphysics of strong emergence and reduction is the fact that questions of integration seem to come *after* questions of individuation (if the latter come at all). Discussing whether an object is physically unified or to some degree physically integrated seem to have been considered (at least by many) as of minor or no importance, possibly because of (non-metaphysical or discipline-specific) parsimonious level exclusivity (e.g. deliberately excluding sub-atomic volatility) or metaphysical assumptions about category or kind (e.g. object *O simply is* physically integrated because it is a member of class C or set S). It seems to me that the more reasonable and cleanly investigative sequence is (1) considering the degree to which the object is integrated or unified *first*, and then (2) considering the status of the thing as an individual, or being able to be individuated (i.e. able to be studied in isolation or apart from the object's environment). Only once it has been individuated can (reducible or irreducible) properties be ascribed to it: hence, to talk about emergence (at least in the metaphysics of emergence) seems to be synonymous with talking about wholes (or at least units of analysis). So, to examine claims in the metaphysics of emergence is to assume not only individuation ('we are studying object *O* and not any non-*O*'), but also integration ('object *O* is, at least at t_1, physically nothing but object *O*').[31]

I believe the first task for us here is to confirm that the metaphysics of emergence needs an object, a unit of analysis, and within the discourse itself, a unified whole. Again, this means that one needs an object to pin properties to—and this object cannot be many things, but one singular thing. Now, emergentism has always been pitted against reductionism—though, as we will see, this dichotomy may have been operating with significant assumptions about 'whole-ness'. It seems to me that they

[31] This last parenthetical statement can be rephrased as follows: *O* is exclusively composed of discrete physical parts, of which we know its totality, and such is *O*, or $O = \{o_1 \ldots o_n\}$. We can individuate *O* qua object iff we have a predicate for '*O equals*', with such a predicate representing a totality and therefore an integrated object.

are initially only opposed in the degree to which the unit of analysis—the thing that someone is observing and investigating—is integrated, and not necessarily what properties it has or even what kinds of units it is composed of. Is it really a singular, possibly unified thing (on one side of the common binary), or is it just a fragmented and/or temporary collection or aggregation of other smaller things (on the other)? As mentioned in the Introduction, modern emergentism is an interjection to the promise of an ultimate, computable, programmable explanation that has continued to fall short of its own expectations. Understanding the (ontologically) reductionist confidence is indeed difficult, though strong and ontological emergentism was *and is* hardly humble in its own right. Let us see how both of these types of perspectives portray *wholes qua units of analysis* and the questions of integration and individuation; for the most part, I will lay out the authors' positions with minimal commentary first, and follow with analysis.

1.1. Reductionism in the present context

Ernest Nagel and his colleagues considered the early emergentists[32] as possibly vitalistic, dualistic, or just trivial, and the former had every confidence in the 'nothing but' argument: all higher-level entities (A's) are nothing but aggregates of lower-level units (B's). But there is much more to reductionism than that, and I would briefly consider three senses of *reduction* that could be distinguished here at the outset. Reductionism as (1) an investigative methodology may be the practice of dividing and sorting, or looking *into* to see what is above or what is 'all' or 'one': a more thorough study of A takes B's into account and holds them in high, if not in the highest, regard. Such a practice may directly or unwittingly be based on assumptions in, or at a minimum have philosophical implications for (2) ontology and metaphysics, where reductionism is the premise that all A's *themselves are* nothing but many B's: to study A *is to study* B's. When these two are blurred together over generations and ripples in a given *épistémè* (e.g. Foucault, 1966), I might argue that reductionism has also become (3) a scientific and popular worldview: because A's are really just B's, and because studying A's is more efficient and thorough when we take B's into account,

[32] Again, the reader is encouraged to read Brian McLaughlin's "The rise and fall of British emergentism" (1992) for an extensive treatment.

we should expect that learning about the world is enhanced when we separate the higher level into lower levels. For our purposes, the most obvious problem with these senses of reduction is that they (seemingly must) presuppose the individuation of both distinct A's and B's, assuming both are kinds of objects that are potential parts as well as subject to property ascription (if we continue to assume that properties can only be ascribed to objects). This leads to, among other outcomes, the explicitly anti-emergentist perspective that the higher level and its properties cannot be studied independently of the parts and properties of the lower level.

The reductionist methodology (1) has been, broadly speaking, enormously successful, and has pushed scientific practice in a myriad of ways to look deeper and further for more fundamental units of analysis and their rules of behaviour. I will not address the socio-cultural reductionist worldview (3), but rather focus on the ontological and metaphysical implications and assumptions (2) typically embedded within this methodology (1). The general, vulgarised amalgamation of reductionism is simple, intuitive, and uncontroversial: *one way* to describe and explain the world around us is to look at what the world is made of, breaking things down into components and studying their interactions. Reducing the properties of a unit of analysis (an object *O*) and its properties ($P_1 \dots P_n$) on the macro-level (*M*) to its parts ($o_1 \dots o_n$) and their properties ($p_1 \dots p_n$) on the micro-level (*m*), as generically depicted in Figure 1.1.1, is really 'step one' of all investigation.[33]

[33] As mentioned earlier, I will often rely on variables or symbols—like O and P_1—to denote the importance of *particular* objects as being things unique to the conditions of the (thought) experiment. There will be a constant emphasis on *irreproducibility*.

Figure 1.1.1: General Reductionist Methodology

M: What are we dealing with? (*O*) → *M*: What is *O* made of? ($o_1 \ldots o_n$)

M: What is *O*'s behaviour?' ($P_1 \ldots P_n$)
M: What is a description of *O*? ($P_1 \ldots P_n$) } *M*: How do $P_1 \ldots P_n$ relate to *O*?

$$O = \begin{cases} m\text{: } p_1 \ldots p_n \\ m\text{: } o_1 \ldots o_n \\ M\text{: } P_1 \ldots P_n \end{cases}$$

At the microphysical level, it might be more property-based,[34] *and less object-based, whilst expressing a sensitivity to or dependency on context or environment:*

M: What is the behaviour ($P_1 \ldots P_n$) of particle/field *O* under condition Y? → *M*: $O_Y = (P_1 \ldots P_n)_Y$

Figure 1.1.1 shows that a general reductionist methodology is ultimately instrumental: it does work, and it does it well at comparatively lower levels of analysis. Being methodological, there are no necessary metaphysical or ontological dimensions; and, empirically, it assumes property ascription will be safe enough to inform the construction of models. The figure above is simply a generalised view on how lower-level science practice can and does work. In philosophy, of course, these issues are not peripheral, but central, and that is my terrain in this thesis: what are the metaphysical implications of *assuming* that *O* is really just $\{o_1 \ldots o_n\}$ (exclusively proper parts) or that $O_Y = \{P_1 \ldots P_n\}_Y$ (condition-dependent, potentially causally-efficacious system-level properties)?

Kim's assertion that philosophy had been under the reign of emergentism since the early 1970s was based on his view that reductionism, vis-à-vis Nagel and Carl Hempel,[35] was fading. This claim is substantiated by the rise of complexity approaches and their associated forms and practices.[36] If a relatively large, dynamic object exhibits

[34] The importance of this kind of property-ontological reduction cannot be overstated, as the reliance on properties at the microphysical level to identify and describe events and behaviour of entities and fields is central to all physics. As Whitehead (1933, 157) observed, "For physics, the thing itself is what it does."
[35] Nagel and Hempel are cited here for their rejection of emergence. Highlights might be Nagel's *The Structure of Science* (1961) with his so-called "bridge laws" for reduction, and Hempel's *Aspects of Scientific Explanation* (1965).
[36] Considering "complexity" as a coherent set of approaches can be at times, like the philosophy of emergence, rather tenuous. Nevertheless, emergence might be considered to join the ranks of self-organisation, co-evolution, chaos, systems, information, actor-networks, autopoeisis, cybernetics, artificial intelligence, and others as emphasising interdisciplinarity, nonlinearity, horizontal relations, and exponential connectivity and dependence.

properties and behaviour that appears to be irreconcilable with an exhaustive knowledge of its constituents (and their properties and behaviour), there arises the question of novelty and how it might be explained. Reductionists in their most benevolent form essentially would claim that the purported novelty is an as-yet unexplained and currently unpredicted (and not unpredict*able*) effect of the combination of particular smaller units in a given arrangement and set of conditions. This, as we will see, has significant parallels with weak and epistemological emergence. Strong and ontological forms of emergence would interject and claim that the novelty represents some unique and irreducible configuration with distinct powers (in the former case) or a fundamentally different kind of stuff apart from what it is made of (in the latter case).

John Searle's five types of reductionism[37] begin with this latter (1) *ontological* form of reduction, which supposes that A's can be reduced to B's if and only if A's are nothing but B's. Two obvious examples are: a) genes are really just collections of DNA molecules, and b) sand heaps are really just collections of particles of sand. (2) *Property-ontological reduction* is a twist on ontological reductionism in that it addresses the actuality of novel properties, but these properties can be deduced from and proportional to their underlying tendencies (as in heat being a mean kinetic energy measurement of molecular movements). (3) *Theoretical reduction* is when laws or tendencies of a reduced theory can be deduced from the laws or tendencies of the reducing theory, such that "the reduced theory is nothing but a special case of the reducing theory." (Searle, 1992, 116) One example might be Kepler's laws of planetary motion being encompassed by Newtonian mechanics. (4) *Logical* (or definitional) *reduction* is a semantics game where descriptors of one unit of analysis can be translated into descriptors of another—such as how a gene is determined by its base-pair sequence. (5) Finally, *causal reduction*, which can relate to properties, concerns what Searle calls the causal powers of a whole as being caused by causal powers of the part. The example he gives shows that a solid object is impenetrable by other solid objects, but this power is only caused by the vibratory movements of molecules in lattice structures. Another three, perhaps better examples are: a) pain qualia caused

[37] Searle, J (1992) "Reductionism and the irreducibility of consciousness" p69-80 in Bedau et al (2006); adapted from chapter five of *The Rediscovery of the Mind* (1992, 111-126)

by neuron firing, b) phenotypic expression 'caused' by (or supervenient on) genotype, or c) our solar system's revolution around the Milky Way being caused by the gravitational pull of Sagittarius A-Star—the supermassive black hole in the centre of the Milky Way. Put together, these styles of reduction prove that the term can be more complicated than the original statement about A's and B's.

Robert Van Gulick[38] asserts that reduction can be understood as a relation between objects in the real-world (what he calls ontological) or between representational items. For him, ontological reduction takes place between: (1) objects in two domains (e.g., minds and brains, or pains and neuron firings); (2) properties (e.g., feeling pain and having neural activation of type N_p, or wanting a cup of coffee and being in a neurofunctional state of type Nf_d); (3) events (e.g., Bill's having a red visual experience and Bill's brain being globally active in a way that includes neural activities of type N_{rve} in his visual cortex as part of its global focus); or (4) processes (e.g., my recalling of the cellist's performance and a sequence of reciprocal neural interactions RNI_c between multiple limbic and cortical areas). Representational reduction takes place between (1) concepts (e.g., links between our first-person concept of phenomenal red and concepts from neuroscience); (2) theories (e.g., links between theories of conscious experience and theories of global brain function); (3) models (e.g., links between models of consciousness and models of reciprocal brain activity); or (4) representational frameworks (e.g., links between the phenomenal first-person descriptive/explanatory framework and third-person neuroscience frameworks). (2001, 3-4) [39]

He goes on to draw attention to five ways in which ontological reduction might be utilised to explain relations between two real-world items: elimination, identity, realisation, supervenience, and composition. *Elimination*, originally from Kemeny and Oppenheim (1956), makes a statement about past and future knowledge, because it recognises cases where "we come to recognise that what we thought were X's are really just Y's." (2001, 4) *Identity*, largely from Place (1956) and Smart (1959),

[38] His "Reduction, emergence and other recent options on the mind/body problem" has proven to be great help in organising thoughts around these issues.
[39] Van Gulick's 'representational reduction' is of minimal interest here, as it does not provide any particularly important distinction from the ontological category in the context of this present chapter.

continues to accept the existence of X's but comes to see that they are identical with Y's. In other words, we see that "two distinct reference routes converge on the same item," and we are not led to eliminate or deny the existence of the prior items. *Realisation*, heavy in mental causation discussions, deals with two different systems that can manifest the same functional property even if it is realised by different structures in the two cases.

The *supervenience* approach offers an escape from composition and identity, though also largely in the philosophy of mind, by attempting to explain the relation between 'normative' properties (moral or aesthetic properties) and their non-normative bases (Davidson, 1970; Kim 1982). One set of X-properties supervenes on another set of Y-properties such that there can be no X-differences without Y-differences. In the normative/non-normative debate, the supervenience slogan would be 'no mental difference without a physical difference.' This topic will be given further treatment in the next section because, as is the case with composition, it can be problematic for both reductionism and emergentism.

Finally, *composition* considers the mereological entirety of the object's constituent parts, and hints at whether the object can maintain certain properties despite the addition or subtraction of some of the parts. This is for me problematic and leaves the door open for other issues, as the author notes, for example:

> To say that (1) a thing is composed entirely of physical parts is not the same as saying that (2) all its parts are entirely physical; to assert the former is to say only that all its parts have physical properties, but the second asserts as well that those parts have only physical properties. (2001, 7, *numbers added*)

To be clear, Van Gulick in (1) is stating that each and every member of the set $\{o_1 \ldots o_n\}$ is entirely physical. For example, a basic desktop computer is composed of a motherboard, a CPU, a memory chip, a network, graphic and sound card, a power supply, a fan, and a case. It is not composed of non-physical parts (e.g. essence, mind, god, soul). Claim (2) says that there is a set $\{o_1 \ldots o_n\}$ where all members are physical objects, and each of these physical objects has properties that are exclusively physical (having mass, composed of elements, reflecting certain kinds of light, capacity for transmitting information, etc.). In other words, each of those objects, such as o_1, has a set of properties $\{p_1 \ldots p_n\}$, and all members of the set $\{p_1 \ldots p_n\}$ are physical. He is right to say that these are two very different claims.

As we will see, this highlights a fundamental problem in the metaphysics of emergence and becomes a major theme of this thesis: should an object be studied via part, or via property, in a discussion of potentially emergent wholes, parts and/or properties? And, when making this choice, what are the implications for individuation? Parts will contend with properties in ontological discussions, and when they do, new epistemological questions arise:

- *(1) Revised*: We know, or have observed, each and every member of the set $\{o_1 \ldots o_n\}$, and they are all physical parts of the whole.

- *(2) Revised*: We have access to all the properties of each and every member of the set $\{o_1 \ldots o_n\}$, and they are all physical properties.

In (1) revised, the claim is that the object's status as a basic desktop computer requires that nothing be added or subtracted to this collection of parts, the properties of which are entirely physical. This means the computer is a totality, and our knowledge of it, or at least of its parts, is complete or lacking nothing. Claim (2) revised states that there exists a complete catalogue of exclusively physical properties of the set $\{o_1 \ldots o_n\}$, and we possess it. This is to say that we have access to all possible properties, and thus empirical knowledge of the object is complete and future observations are irrelevant. The former claim is bold, but not nearly as bold as the latter. And, as we move along, it prove increasingly more difficult to untangle the first set of essentially ontological claims with the second, revised set of essentially epistemological claims. It is the entanglement of parts and properties—and ultimately of ontology and epistemology—that provides for many of the flaws shared by both kinds of strong claims. More is to come on all of this beginning in the next section.

1.2. Main doctrines of emergentism

Especially in the forms mentioned earlier, reductionism has seemed to be inadequate for explaining many cases of higher-level complex phenomena. The as-yet inadequacies of cognitive science are the dominant example, but consider the unsatisfactory nature of an explanation of network behaviour in the 21st century exclusively in terms of binary code. It would seem a rather impoverished idea to think that we could understand electronic communication trends based on the flow of 'zeroes and ones', when the code itself is composed whimsically, relationally, and spontaneously as it is being sent. Whilst the text of the message is physically composed of coded segments, a print-out could only be translated into the corresponding linguistic symbols. We might be able to re-assemble the message into a coherent, declarative sentence, but we would be without context and connotation, unable to distinguish between opinion and fact, unable to determine the relations between communicants, and unable to specify environmental events which may have led to and/or affected the transmission. Hence the message has physical, micro-level components which alone cannot describe, define, compose, explain or predict the 'fullness' of the communication.[40]

Complexity movements hoped to at least aid the partial explanation provided by the standard forms of reductionism. Most of the descriptions of emergence though are based on their antithesis from one style of reduction or another, and thus they end up having common terminology. Before a series of various taxonomies and descriptors, an overview of the ontological positions is necessary.

1.2.1. Ontology

As Phillip Clayton[41] and others have noted, any discussion of (strong) emergence presupposes ontological monism because of emergence's qualities of novelty,

[40] Extreme reductionist cynics may try to separate the code from meaning or semantics, but that would presuppose the code or its fundamental elements as being isolated units of analysis and possibly even universals.
[41] Clayton states this directly in "Conceptual foundations of emergence theory" (2002, 2), and this can also be found in Ellis, Kim, Anderson, and a host of others.

unpredictability, and irreducibility. It is precisely because the ontology is monist that the novel properties are significant. If an object is made of only one kind of stuff, how could these properties emerge? This is a far more interesting question than: if the object has at least two different ontological categories, the, for example, non-material category being unobservable or unrecognisable, how could the properties emerge? If the ontology was already pluralist before the claim for an emergent property or object, there would be less motivation to describe *and* explain it as it would not necessarily be unique.[42]

Weak, descriptive, epistemological emergence cares not for ontological categories, as it does not require any presupposition on which or how many different types of stuff there are in the cosmos. She who will become an ontological emergentist though begins the investigation into alleged emergent phenomena holding monism to be true, for the surprise of novelty would not present itself otherwise. Before the claim of novelty, she must at least grant an ontological category to the lower-level objects recognised by physics (physicalism). Clayton fails to explain his claim that there could be a (tenable) monism outside of physicalism before the claim of ontological emergence. If one believed in only one kind of stuff—before claiming that a network or person has emerged ontologically—what kind of stuff would that be other than physicalist[43] bits? To denounce *all* ontological claims made by physics is to remove one's self from the discussion, so before novelty there must be a physical base from which it would make sense.

During and after the claim, she may come to hold one of the following positions:

- *1) All things are still made of one kind of stuff (monism), but some things must be said to configure themselves such that they must be more than mere aggregates of the particles*

[42] There is an important distinction here between object and property, because a novel property in some cases may not prompt the claim for a novel object (a new category of existence beyond that which is recognized by physics). Also, with this phrasing it is easier to see why ontological pluralism is not necessarily interested in ontological emergence *per se*.

[43] Henceforth, physicalism will be understood as considering that which is recognised by contemporary physics as ontologically primary, though not including a singular ontological primitive (e.g. superstring). For some authors, physicalism is misconstrued as being necessarily ontologically reductionist. This is an error, as we will continue to see. 'Physicalist' is preferred over 'materialist' to emphasise the contributions made by physics. However, I prefer a new physicalism in order to emphasise not a bottomed-out unit or ontological primitive, but rather the lowest reasonably possible level of analysis to aid in description and explanation. Nevertheless, I will simply continue using 'physicalism' in its *not-necessarily-ontologically-reductionist* variety.

and forces identified by physics (or standard physicalism). System-level properties seem to be more than just additive properties (e.g. mass, colour, shape).

This is also called non-reductive physicalism (NRP),[44] and it can collapse into weak emergence because it does not necessarily grant a new ontological status to the purported novelty. It also becomes more complicated if the claim is for novel properties and not explicitly for new entities or new causal powers—this will be treated more extensively in the next section. If this timid NRP, which more often than not treats novel properties, is not quite aggressive enough in dealing with the alleged unpredictability and/or irreducibility, the more radical choice becomes:

- *2) Facing strong evidence of at least unique causal powers, physicalism cannot be said to exhaustively categorise the ontological status of the higher or macro-level object in question.*

With this second position, the emergentist must at least be holding a dualist ontology (such as when one is claiming that the mind is of another category from the neurons, etc. that comprise it) and possibly even a pluralist ontology (e.g. Dupré, 1995) where there are many different kinds of things having unique, unpredictable, irreducible properties that probably also have the ability to causally affect their own (presumably physical) constituents.

The NRP ontology remains the default position when one is unwilling to go to the radical position of (2) but is forced to admit some absence of explanation in light of, for example, the failure of the standard models of the natural sciences. It is the equivalent of a metaphysician saying, “I am uncertain about how to explain the existence and/or behaviour of X, so I will suspend such judgment; yet, I will not proceed with any further investigation in at least physicalist terms.” This position seems to be a temporary stop-gap, one which sits idly whilst natural science practice finds time to do what is for philosophy the necessary synthesis.[45] I am not convinced by this approach, and for two reasons. First, I do not think that most metaphysicians

---

[44] In other writings, Kim has expressed his suspicions that non-reductive physicalism is untenable (*Physicalism, or Something Near Enough*, 2005). Again, the majority of his arguments are in based in the philosophy of mind and mental causation.

[45] I use this language to recognise the plethora of philosophical answers and subsequent new questions that have derived from science practitioners themselves, and not from within philosophy—the presumed domain for such synthesis.

have gone far enough in adopting a consilient[46] framework, where there is an active engagement with and synthesis of the full range of mainstream and anomalous natural sciences. Secondly, what is more relevant here is that it should be unsettling to be faced with these two choices above, considering that we do have novelty *of some kind* everywhere—even if it is derived merely from weak system-level property ascription (e.g. shape). It is possible, I argue, to make sense of the novelty without being forced to take one of these less-than-satisfactory positions. It is in the reconstruction of the *whole* (vis-à-vis the unit of analysis) I believe—within a responsible and more open physicalist framework—that this novelty might begin to make sense.

1.2.2. Summarising weak emergence

Any talk of emergence is more significant when it makes a claim for its strong, ontological form. Weak, epistemological emergence is not the target of this present work, but it might help us understand the motivations for claiming ontological novelty. It is my argument, stretched out over these chapters, that novel properties might happen to be unpredicted—and may only seem to be unpredictable—because of a self-imposed ontological restriction concerning the nature of the whole. If this is the case, as I hope to gradually show, property emergence—the exposure and identification of novel properties—might be the first step in acknowledging these limitations. A full re-examination of the possibilities and applications of weak emergence might end up being profitable after the present work is concluded. For now though, I will simply provide a quick overview to flesh out some of the key aspects, because the recognition of novel properties is a necessary precursor to ontological claims. One cannot claim the existence of a new kind of thing without having observable and/or measurable reasons for doing so—assuming one is at least starting with a physicalist position.

[46] This is used in the sense of E.O. Wilson from his monograph of the same name (1999). It is meant to inspire an unlimited pandisciplinary sourcing, where analysis of all natural sciences can not only challenge the standard models but piece seemingly irreconcilable data together to make greater contributions to our understanding of the world.

First, it is undoubtedly one of the most useful tools in modern science[47] (even if not referred to as a separate methodology by practitioners) as it helps to expose—though not explain—new phenomena and properties of aggregate behaviour. Practicing scientists in various fields often exclaim that they dread 'emergent' properties because their presence indicates another loose thread to tie up in their data. It aims to recognise epistemic novelty and nomological unpredictability. Where there are cases in which an exhaustive knowledge of the (physical) parts, as aided by standard models, fails to predict, explain, and maybe even identify properties and behaviour of higher-level entities and wholes, the area of enquiry is flagged as being incomplete. Nagel in *The Structure of Science* (1961) would approve of this notion, as he thought emergence was in any case simply a temporary (or situational) confession of ignorance. There is less of a need in these weak cases to pinpoint causal relations (which might inspire claims of new *things*), as explaining novel or unexpected observations in causally-contingent terms seems to at least temporarily suffice. The example of a traffic jam has been given to illustrate the fact that, although we know the jam is a description of the aggregate behaviour of many individual cars in a roughly-defined, densely-populated space, it is interesting to see how the jam has surprisingly unpredictable properties as it develops.

As Bedau has described it, weak emergence is "metaphysically innocent" (1999, 376) and demonstrative of "explanatory incompressibility". (2010, 51) A cellular automaton (CA) is composed of the same simple units ('cells'), but their aggregated growth over time is not always predictable. The figures below show how even with a simple beginning behavioural rule, with no added environmental effects, behaviour of a simulated closed system can exhibit totally random patterns and behaviour. One of the most fascinating aspects of the CA is the fact that this unpredictability, which is still prevalent over millions of simulated 'time-steps', is completely isolated from other forces: this is an impossibility in nature, and therefore weakly emergent patterns revealed by simulated closed systems should tell us how important the micro-level detail and context sensitivity really are.[48]

[47] It was via weak emergence in cellular automata that one can observe complex patterns arising out of simple initial rules. Conway's 'Game of Life', Holland's complex adaptive systems, and Stephen Wolfram's *A New Kind of Science* (2002) are primary sources for this issue.

[48] Another key aspect of CA studies that, to my knowledge, has been overlooked, is the fact that, for every time-step, little attention is paid to the possibility of a new time-step. Its progression from one to

Figure 1.2.2.1: Cellular Automata Item A

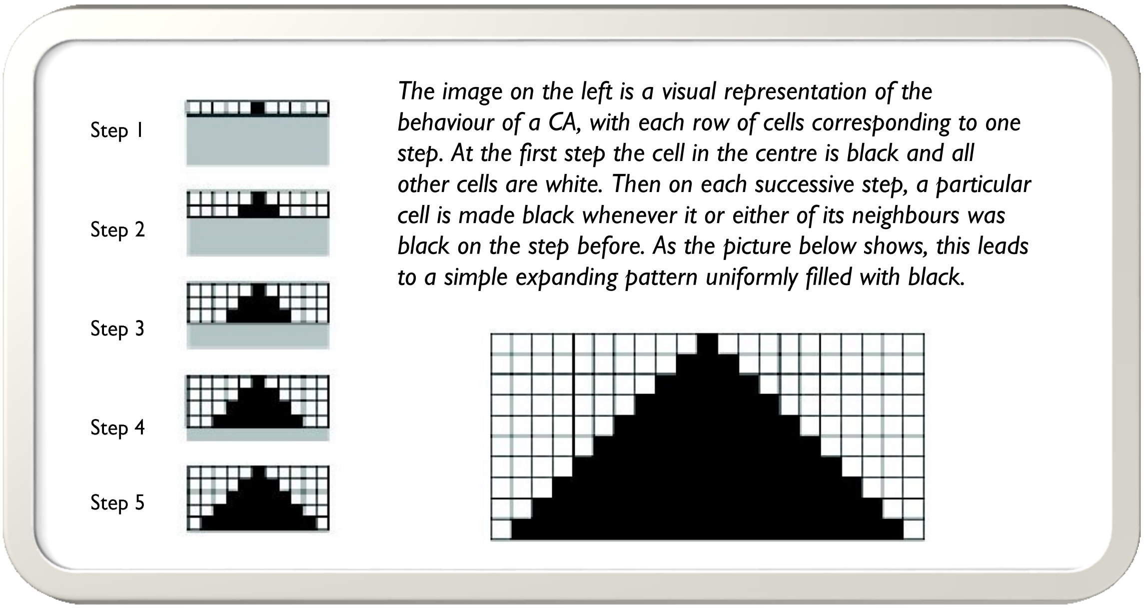


Figure 1.2.1 (top, p33) *Cellular Automata Item A* Images and text © Wolfram Media
Wolfram, S (2002) *A New Kind of Science* Wolfram Media, New York (p24)

Figure 1.2.2.2: Cellular Automata Item B

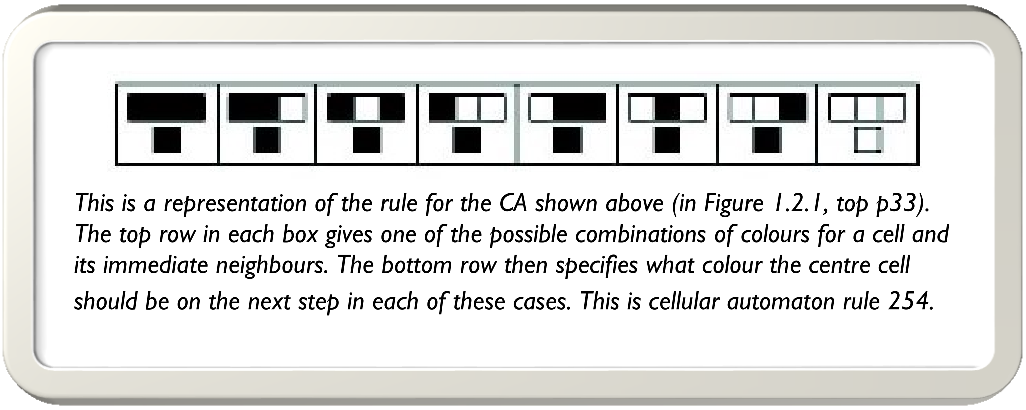
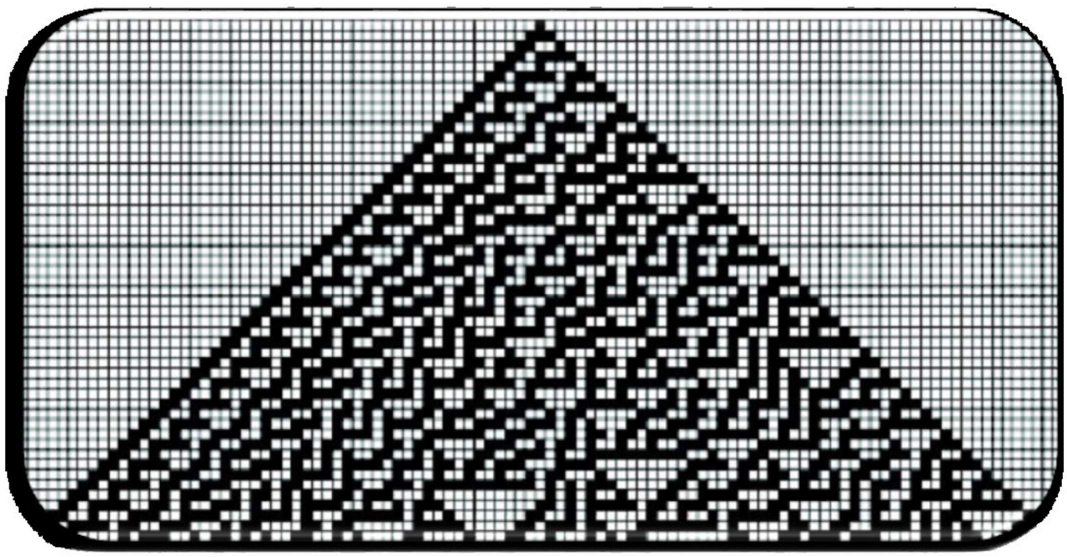


Figure 1.2.2 (bottom, p33) *Cellular Automata Item B* Images and text © Wolfram Media
From Wolfram, S (2002) *A New Kind of Science* Wolfram Media, New York (p24)

another is assumed, and so commentaries and explanations of the unpredictability of CA simulations simply focus on the developing pattern. It is, as I see it, of *overwhelming* importance that new time-steps require not just new time but new space of which the new cells occupy. Why should we assume that the ability of new cells to appear should be taken for granted? Even when using a singular, initial programme (or rule), there still must exist the *environmental capacity* for the programme to run. Even in simulation, the environment, or local physical conditions, can never be ignored in description and explanation.

Figure 1.2.2.3: Cellular Automata Item C



This is a CA with a simple rule that generates a pattern which seems in many respects random. The rule used is of the same ***type*** *as in the previous examples, and the CA is again started from a single black cell. But now the pattern that is obtained is highly complex, and shows almost no overall regularity. The picture is our first example of the fundamental phenomena that even with simple underlying rules and simple initial conditions, it is possible to produce behaviour of great complexity. This is rule 30.*

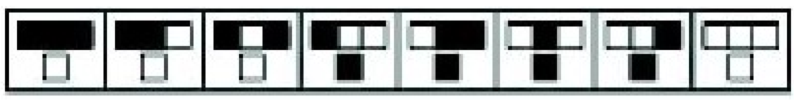
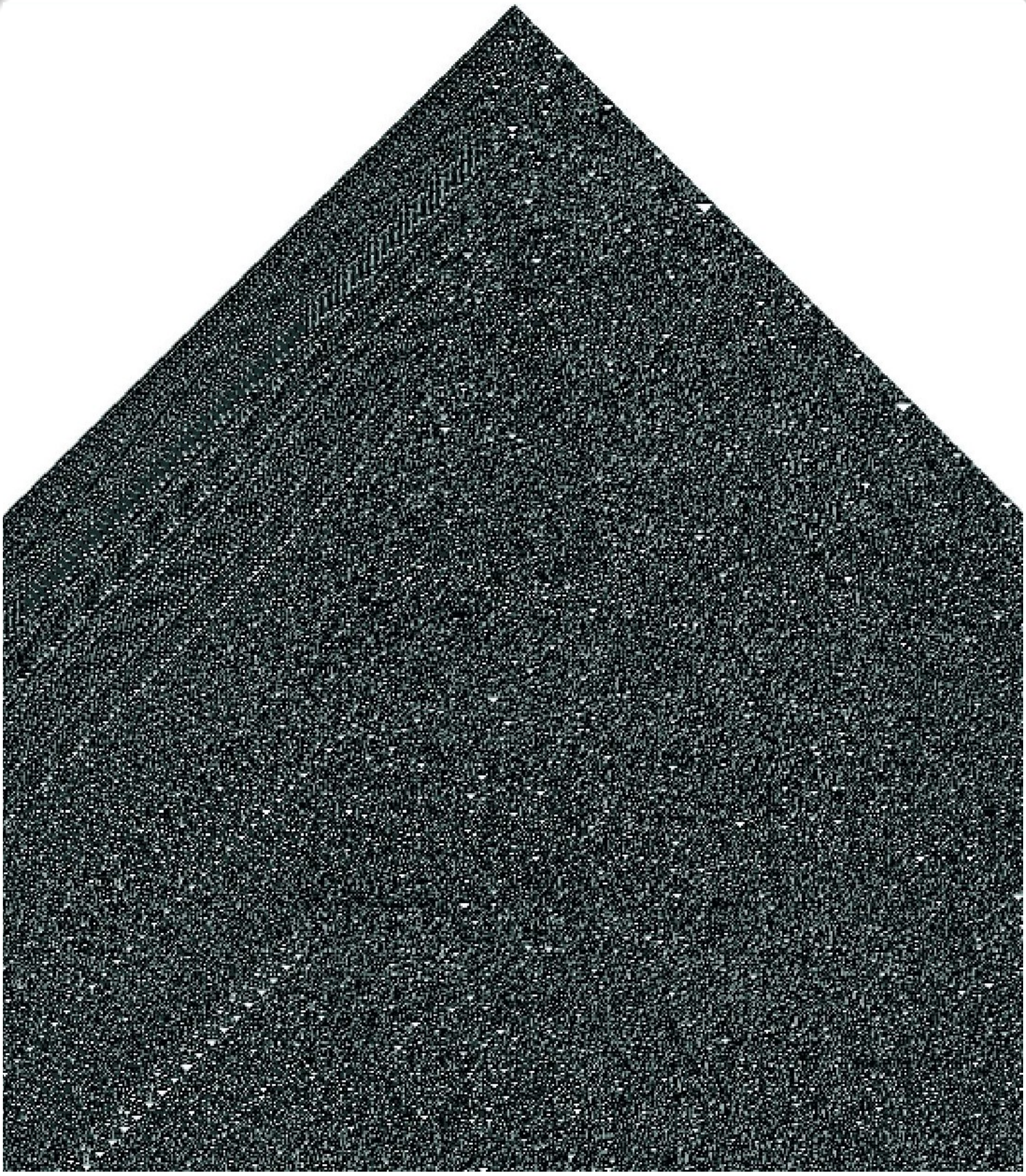


Figure 1.2.3 (p34) *Cellular Automata Item C* Images and text © Wolfram Media
From Wolfram, S (2002) *A New Kind of Science* Wolfram Media, New York (p27)

Figure 1.2.2.4: Cellular Automata Item D



Fifteen-hundred steps of rule 30 evolution. Some regularities are evident, particularly on the left of the image. But even after all these steps there are no signs of overall regularity—and indeed even continuing for a million steps many aspects of the pattern obtained seemed perfectly random according to standard mathematical and statistical tests. The picture here shows a total of just under two million individual cells.

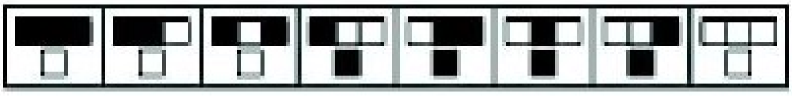


Figure 1.2.4 (p35) *Cellular Automata Item D* Images and text © Wolfram Media
From Wolfram, S (2002) *A New Kind of Science* Wolfram Media, New York (p29)

For Bedau, there are two main features of weak emergence:

> (1) Emergent phenomena are somehow constituted by, and generated from, underlying processes.
> (2) Emergent phenomena are somehow autonomous from underlying processes. (1997, 375)

One can see how the notion in (1) relates to the CA. There is no alternative statement on the ontological composition of the CA, because the individual elements are all of the same category. The emergent phenomena—the patterns—are clearly constituted by the same cells throughout. With (2), the patterns might be said to be autonomous from the generating rule, but this is not terribly insightful in ontological terms. A pattern might 'emerge' or be generated anywhere in the world, but its autonomy can always be disputed simply because it is difficult to claim that a *description* has causal powers or a disassociation from what is comprised of. Hence we are drawn back to the first proposition in §2.1 which states that although we admit to compositional uniformity, the forms that these units take sometimes defies predictability. This kind of NRP says nothing about ontological status, but flags a gap in explanation based on a failure to acknowledge or incorporate the context (or the environment) in which it is found. Weak emergence, therefore, cannot be the source of the contentions around the emergentist corpus, and it is not the *problematique* which is being treated here in these chapters.

Ontological and strong emergence, the elements of which that are being distilled from a range of authors here, is the focus, because it is where emergence fails to reconcile physicalism and alleged ontological and etiological novelty. Before I begin the survey, it is important to realise that some elements of weak and strong are woven together for many of these authors, and only later in this section will I present the noteworthy strong claims. Ontological claims (and not necessarily that which we have up until this point been calling 'strong' claims) must invariably discuss a new category of existence, and there are numerous pathways for scrutinising them. Strong claims, which I will distinguish from the ontological in the coming paragraphs, typically address the causal capacity of either novel objects or novel properties.

I choose in this thesis to look at these two types of claims via *wholeness*: I hope to show that some reductionist and emergentist ontological and causal claims may be based on tenuous, outdated, or omitted (and/or assumed) notions of the whole that appear to overlook contributions from contemporary mainstream science, various other complexity sciences, and other explanatory analytic options. It will become clearer as we progress that what I consider to be errors and/or omissions in wholeness originate in errors of individuation and property ascription. Let us look now to see what these authors are discussing, and focus in on the claims made concerning the nature of (the properties of) the whole.

1.2.3. Van Gulick's three kinds

As a sort of introduction to emergence, Van Gulick's survey article can help in identifying three broad kinds. The unimpressive (but perhaps intellectually thorough) *specific value emergence* (1) includes cases where the whole and the parts share features of the same kind, but have different sub-types or values of that kind. His example is that of a one-kilogramme bronze statue, whose molecules clearly do not each have a mass of one kilogramme. These kinds of properties might be considered as purely additive.[49] This first kind was probably included simply because the whole (the statue) has such additive properties that the molecules do not, because the molecules' individual masses are not each one kilogramme. The second, or *modest kind of emergence* (2), is where the whole has features that are different in kind than those of its parts. The example here is of a purple cloth whose molecules could not be said to individually be purple. Perhaps a better example might be that of a mouse, which might be alive even if none of its subcellular parts could be said to be alive. As Kim (2006b, 549) notes in reference to Van Gulick, it is unclear how these properties can be said to "emerge" from the properties of the constituents.

The statue's additive properties are not only weakly but perhaps sluggishly emergent, and rest entirely on a) the notion that the statue is whole, even if just in an artistic sense; and b) that the properties of mass are the only properties to be considered in this ascription of *wholeness*. The purple cloth's holistic property of 'purpleness' is

[49] Additive properties, as will be shown in later chapters, are contrasted with interactive properties: this distinction is drawn in the work of Joseph Earley (e.g. 2005).

attributed to the aggregated refracting properties of the molecules which comprise it: this means the cloth would not be purple without light. Imagine a ‘purple cloth’ in a pitch-black room: is the purple cloth still purple? Is purpleness an intrinsic property of the cloth? The purpleness might be attributed to the molecules and atoms, but the purpleness cannot be intrinsic to the cloth, because the cloth only appears to be purple under certain lighting (amongst other) conditions.[50] Hence the existence of the purpleness property is dependent on environmental factors. The claim also presupposes that the whole (the cloth) is an individuated object and wholly distinct from its environment. Thus the purpleness is only weakly emergent if we assume the cloth is static and conditions are irrelevant or remain within a normal range.[51]

Van Gulick’s third and *radical kind of emergence* (3) has two criteria. First, the emergent property is different in kind from those had by its parts {the same as (2)}. Second, a whole’s emergent property is of a kind whose nature and kind is not necessitated by the features of its parts, their mode of combination, and the law-like regularities governing the features of the parts. This second condition is a crucial distinction from the other kinds, because it means that the emergent property is to some degree disassociated from its micro-structural properties. The author believes that it is unlikely that any cases of this third kind actually exist, and that they would violate his idea of atomistic physicalism. Much more importantly though, emergent properties of this third kind cannot be said to supervene on the lower-level properties, and this would indicate a violation of microdeterminism or mereological supervenience.

He does not give any examples of this because he does not believe any exist. But, even if he were looking for them, he has arranged the quest such that he will never find what he does not want to find. That is to say: when phrased like this, the deck is stacked against strong emergence. He expects to never find static, yet emergent wholes—and is right to think so because such things have only existed in the minds

[50] The cloth would appear to be more white than purple if the room was flooded with bright light, as the small amount of purple light refracted by the cloth would be overwhelmed by the white light around it. Furthermore, the purpleness of individual molecules would still need to be detectable a) under precise lighting conditions, and b) using sophisticated equipment and not the naked eye.
[51] Normal here implies standard earthbound conditions in temperature, pressure, gravity, climate, etc. all within a suitable range for mammalian-like survival, detection, and computation.

of philosophers. Unfortunately, he does not admit to this, nor do those who advance similar positions. The tenuous construction of wholeness is overlooked, and the discussions carry on speaking of the mode of combination and law-like regularities governing the features of the parts of the whole in question. By his own admission he is basing his predication on pitting atomistic physicalism against “radical” (already a daunting term) ontological or strong emergence. The physicalist element is expected, because we must be expected to explain physical objects/events in physical terms. The atomistic element is the static, set-theory-inspired element that bottoms out with distinct, individuated, instantiated objects. It entails pure, ontological reduction, and I argue that this false dichotomy is crippling both the reductionist side and the strongly emergent side.

There can be no bottom unit with an updated physicalism, and, by physicists’ own admission, there is no resolved, ontologically-independent ultimate entity supported by any observational evidence.[52] Atoms can be used to describe the lowest level of recognisable and distinguishable entities, but our explanations can never bottom-out with them. If we choose to have as our unit of analysis something which is comprised of other things we have not accounted for, such as sub-atomic particles, how can we claim to know the precise nature of our very own unit of analysis? Furthermore, we know that as the level of composition moves down, the sensitivity of the units on that level towards the environment increases. This is to say that tracking and identifying the behaviour of very small units becomes increasingly more difficult as they are observationally more sensitive to microphysical changes typically unobserved (or deemed insignificant) on higher levels.[53] Hence, if we want to argue that a claim for radical ontological or strong emergence has violated atomistic physicalism, we would need to have amassed a large amount of observational as well as environmental data.

[52] Are they strings, as in string theory? If so, what would determine the nature of the strings? The goal for many physicists is to find that bottomed-out unit, and yet the first question is again, “what is it made of?” Infinite regression may be inevitable, but it is what asymptotic mathematics and investigation is all about: getting closer, more refined, but never *there*, or at an end. The success of QM has been described in one sense as being able to measure the distance from New York to Los Angeles to a precision equal to the width of a human hair; as such, science zooms *in* but never *to*.

[53] In the wake of a recent experiment (CERN, 2011) that initially seemed to challenge mass-energy equivalence (the Einsteinian rule that no matter or information could travel faster than the speed of light), one of the proposed explanations for the strange data was environmental interference in the measurement process. Sure enough, after repeated testing, multiple groups found the original results to be insignificant or invalid. (e.g. CERN, 2012)

I will return to these ideas in Chapter III, but for now the contributions of Van Gulick on this radical kind must be cautiously guarded.

1.2.4. Kim's five claims

For Kim and others, with their version of a physicalist foundation, macro-level properties can be considered emergent if they are both non-supervenient and irreducible. This means that the properties of the whole must, to fit this description, a) not depend on the properties of the parts for their own existence, and b) not be recognisable based purely on a collection of the properties of the parts. He states that these conditions come together to exhibit downward causation (a strong type of macrocausation) which is "the *raison d'être* of emergence."[54] Before arriving at that conclusion however, and before going into detail on these terms, let us look at how he arrives at this position. In "Making Sense of Emergence" (1999a), among the most frequently-cited articles in recent philosophical discussions of emergence, he identifies five features that emergent phenomena or properties might need to have. These features attempt to take into account the entire history of emergentism beginning with Mill, Lewes, Alexander, Broad, and Morgan.

The first is the (1) *emergence of complex higher-level entities*. Systems with a higher-level complexity emerge from the coming together of lower-level entities in new structural configurations (or the new 'relatedness' of these entities). This is nothing surprising, and it is also true for reduction, for it is also based on a fundamentally physicalist position where matter and energy behave according to physical laws. The difference with emergent phenomena for Kim is the idea that non-physical (or *normative*, because he is focusing heavily on philosophy of mind) properties exist outside the domain of these laws, and the remaining claims try to explain why. However, the author has not stated his position on the nature of physical laws—a very much disputed terrain[55]—and he also commits an error he will repeat later. Being

[54] This quote comes from Kim's later article "Emergence: core ideas and issues" (2006b, 548).
[55] Where the 'laws of nature' come from is obviously a classic problem in the philosophy of science. I subscribe to a similar position to Paul Davies (e.g. 2010, *Information and the Nature of Reality* Cambridge), who supposes (though he postulates with quantum information theory) that what we consider as laws or constants are never really static and fixed. In essence, Platonic mathematical truths

selective about which levels to include in the analysis of wholes severely hinders one's argument: one cannot consider quantum and atomic information in one place, and then ignore it in another. In any case, we will proceed with his taxonomy.

The second feature is the (2) *emergence of higher-level properties*, where all properties of higher-level entities arise out of the properties and relations that characterise their constituent parts. Some properties of these higher, complex systems are "emergent", and the rest merely resultant. Kim notes that the expression "arise out of" could be exchanged with "supervene on" or "are consequential upon". He goes on to say that the key point to this second claim is that when the parts are configured in a certain way, the system will necessarily exhibit certain higher-level properties. It is important that, in this feature of his, both emergent *and* resultant properties supervene on (or arise out of) their micro-structural (micro-based) properties. In other words, this second feature (which presumably incorporates the first) permits a kind of NRP: something about the higher-level distinguishes itself from the lower, but not at the expense of the ontological status of the lower. After these first two claims, Kim has not presented us with anything terribly different from the core notions of weak emergence as described above; basically he is saying that the system-level can be special, but not too special. I suppose one example might be an inflated football, where there is some organisation (i.e. the ball does not explode, most of its physical particles do not dissipate, it maintains a shape, a force applied at one point drives the whole object in a given direction, etc.), but it seems to lack autonomy, and most of the object's behaviour is explained by simple Newtonian mechanics.

The problem of distinguishing between the emergent-resultant properties is dealt with, it seems, by predictability and explainability. The third feature then is the (3) *unpredictability of emergent properties*, which simply states that emergent properties are not predictable from exhaustive information concerning their basal conditions. Resultant properties can be expected based on lower-level information. The fourth feature is the (4) *unexplainability and irreducibility of emergent properties*, which, also

are ultimately a futile aspiration. He considers—correctly in my opinion—that local interactions determine what we call global laws or global models, and also considers the possibility that it is entirely unhelpful to assume that global laws determine local interactions. Generating models, or a map of behaviours (properties), is perhaps as far as we should go towards establishing scientific truth (and, indeed, this is the case for most, even unwittingly so).

unlike resultant properties, cannot be explained by or reduced to lower level information. These two claims can be paired together in light of their connection to reduction. To predict or expect properties at the system-level is to declare what a set of properties will be based on the memory of similar systems with similar (or identical) constituent parts. To explain or reduce properties at the system level is to declare those forces (perhaps from the bottom-up) that had seemingly caused the system-level properties to be what they are either by process, function or composition. I remind the reader of the example from earlier of electronic communication: some systems or objects seem to exhibit properties that the physical and microphysical components cannot entirely account for.

Even from the third and fourth features, the stakes are being raised as to the necessitation and potency of higher-level properties. The last feature is the most tenuous, and it concerns the (5) *causal efficacy of the emergents*. Here emergent properties have their own novel, causal powers that are inexplicable by and irreducible to their lower-level properties. The argument for macro-level causation must feature one or both of these: a) the ability of the system qua system to have unique causal powers; and b) the ability of the system to have unique causal powers directed towards its own constituents. The analysis should stop here, as extensive treatment of downward causation will be presented in Chapter III. What we must take away from Kim here is that (3) and (4) can easily recede back into weak emergence, and that they are most interested in properties. The claim that the properties are irreducible does not demand a new category of stuff, but it does require an ascription of wholeness to the object holding this property. The fifth claim does this also, and unlike (3) and (4), it leads one into the realm of ontological or strong novelty in light of the alleged novel causal powers of its individuated system level.

1.2.5. Ellis and Deacon's five levels

Different levels of emergence have been suggested by Terrence Deacon in *The Symbolic Species* (1997), and George Ellis in "On the nature of emergent reality" (2006) has attempted to reformulate Deacon's levels with a slightly different characterisation. These levels are progressive and hierarchical, and demonstrate a similar sense of progression as Kim's claims and Van Gulick's kinds. As they progress, the stakes and the importance of concrete examples increase.

Level-one emergence involves bottom-up action leading to higher-level generic properties but not to higher-level complex structures or functions. Examples include the determination of generic properties of gases, liquids, and solids, such that gas laws, conductivity, heat capacity can be determined or approximated. This kind of emergence leads to coherent upper level action, and reduction is in principle possible. For Ellis though, static, ontological reduction fails in many cases because of, for example:

> (1) the inability to derive in this way the full complexity of behaviour of substances as simple as water
> (2) the arrow of time/entropy problem is unresolved
> (3) quantum measurement issues are unresolved
> (4) divergences and incorrect predictions of the value of the cosmological constant mean we do not properly understand quantum field theory
> (5) R. B. Laughlin's claim that all elementary particle properties may be emergent
> (2006, 93) [56]

Level-two emergence concerns bottom-up action plus boundary conditions that lead to interesting higher level structures not directly implied by the boundary conditions. Ellis' examples include (i) sand piles, (ii) the reaction diffusion equation, (iii) magnetic domains, (iv) convection patterns, (v) cellular automata, (vi) gravitational structure formational in the expanding universe, and (vii) inorganic and organic molecules. The system can be said to have increased in complexity, and it leads to the emergence of structures that are not reducible. Despite the immense combinatorial possibilities[57] they are not truly complex, as they do not have "the key element of goal-seeking that characterises living systems." (96)

Level-three emergence involves bottom-up action in highly structured systems that leads to a) the existence of feedback control systems at various levels as well as to b) coordinated responses that allow meaningful top-down action. This then also leads to coherent non-reducible upper level action directed by implicit in-built goals. Ellis writes:

---

[56] With some of the description of Ellis' (Deacon's) levels, references have originated in the Ellis text, such as (in this instance): Zeh, H D (1989); Halliwell et al (1994); Penrose, R (1989); Laughlin, R.B. (1999)
[57] Scott, A (1995)

> Thus these systems enable an element of teleonomy[58]–of goal-seeking representing an effective physical effect of information. However the inherited goals guiding these feedback systems are independent of individual life history, being pre-determined by the evolutionary history of the species - no learning occurs. This allows adaptive behaviour, but based on pre-set rules. Examples are processes in all living cells and in plants. This information-based functioning starts at the supra-molecular level. (97) [59]

Level-four emergence, in addition to level-three as above, concerns the existence of feedback control systems directed by explicit goals related to memory, such that the 'mind' has been influenced by specific events in the individual's history. Learning occurs based on individual experience and some form of stored memory, allowing adaptive behaviour responding to historical events. This fourth level, in a somewhat bizarre manner, shifts to what is apparently exclusive to the mental realm, something extremely rare on a cosmic scale. However, the levels might be more widely applicable if references to the mind were replaced with references to computation (as is the case with the next section, §2.6).

Level-five emergence, in addition to level-four as above, involves some goals that are explicitly expressed in language systems and/or are determined by symbolic understanding or complex modelling of the physical and social environment. In addition to individual 'consciousness', the capacity to handle symbolic systems with both syntax and semantics (Deacon), presumably arises in conjunction with the capacity for self-conscious reflection. Integral to reflection is the feature of *distributed consciousness*, with the development of brains and culture occurring in interaction with each other.[60] This final level is the most ambitious, and perhaps, vague, but nevertheless demonstrates the escalating complexity and its strange effects. The exponential development of 'emergence' in these five levels might be an accurate representation of the contrast between conscious beings and the cells which compose them in terms of interconnectivity and functional capacity.

---

[58] This is a contentious phrase it seems, considering the (Aristotelian) teleological implications.
[59] Harold, Franklin (2001); Bidwell, R G S (1979); Lehn, Jean Marie (1995)
[60] Berger, P.L. and T Luckmann (1967); also Merlin D (1991) and (2001)

The last two level descriptions, especially, relate to an increasing system-level complexity with advancing powers and greater autonomy. As levels four and five identify a higher-level progression—from inferior cognitive capacity (to sentimentality) to advanced memory storage, to verbal communication with other beings, and finally even to self-regulation—in some ways reminds one of this passage from *The Descent of Man*:

> Firstly, the social instincts lead an animal to take pleasure in the society of its fellows […]. Secondly, *as soon as the mental faculties had become highly developed*, images of all past actions and motives would be incessantly passing through the brain of each individual […]. Thirdly, *after the power of language had been acquired*, and the wishes of the community could be expressed […]. Lastly, habit in the individual would ultimately play a very important part in *guiding the conduct of each member* […]. (Chapter III, 71)

Ellis and Deacon's levels are seeking a teleonomical feature of alleged emergent wholes, and as such, they would not be interested in hearing about strongly emergent properties outside of cybernetics, cognition, and developmental biology. The mention of boundary conditions—and the presupposition that insufficient matter and energy are exchanged with the environment, and thus having insufficient effect—is misleading considering the role that the environment clearly has in these cells and systems. If our analysis is limited in scope and discipline, we would seek to explain biological systems, for example, purely in biological terms. But, since the claim here is for the existence of real novelty, either in properties (One through Five) or in stuff ('Five'), on multiple levels of reality using different units of analysis, one should expect, for example, quantum (endogenous) and environmental-global (exogenous) effects. How could we possibly claim ontological novelty if the systems we presume to exhibit this novelty are whole and static and are disassociated from microphysical changes? We are again forced, I believe, to remain sceptical of the ascription of system-level novel properties when the system itself was individuated and often instantiated for parsimonious analysis.

1.2.6. Assad and Packard's three types

The authors of "Emergent colonization in an artificial ecology" (1992) may not normally be included in metaphysical exchanges about emergence, largely because their interests—even in the paper itself—lie in artificial intelligence, computer science and cybernetics. The section of the paper that is useful in our present context is also quite short and vague, but it presents a unique perspective on system-level analysis. The first part of this section identifies four types of emergent levels, not wholly unlike the levels mentioned previously. First, non-emergent levels are where behaviour is immediately deducible upon inspection of the specification or rules generating it. Second, weakly emergent levels are where behaviour is deducible in hindsight from the specification *after* observing the behaviour. Third, strongly emergent levels are where behaviour is deducible in theory, but its elucidation is prohibitively difficult. Finally, maximally emergent levels are where behaviour is impossible to deduce from the specification. But, the more significant contribution comes immediately after this. They ask a simple question that is still unresolved: *what emerges?* Three options are provided in a "hierarchy of necessity": (1) structure, (2) computation and (3) functionality. (1992, 3-4)

The emergence of patterned structure (1) exists either in what the authors call "space-time configurations", or in "more abstract symbolic spaces". The reader here will see in the next chapter that this has some similarities with classical mereology (as well as with Van Gulick on reduction), which deals largely with three-dimensional, spatially-extended objects (extensional) and to a lesser degree with mental or representational objects (intensional). Also, 'structure' here might imply that the system has begun to organise itself based on maximal efficiency and stability (and minimal energy loss). Even if there is some demonstrable computation which has been automatically implemented in the formation of a structure, the emergence of computational processing (2) is more impressive because it exhibits computation by the system qua system and not at the micro-level. This resonates with macrocausation as the higher-level entity has newfound downwardly-directed causal powers. Functionality (3) emerges when system-level actions are functional or beneficial to the lower-level components as well as to the system itself. The authors claim that the emergence of functionality is exemplified often as being a "subset of the examples of computational emergence where computation performs a function" for the lower-level constituents. (1992, 4)

Assad and Packard briefly, and temptingly, mention that there is a hierarchy of necessity where the emergence of functionality requires the emergence of computation—which in turn requires the emergence of structure. Conversely however, as they point out, the emergence of structure does not necessarily imply the emergence of computation, which does not necessarily imply the emergence of functionality. It seems this "SCF" model represents a tidy way to view, in particular, biological emergence, as one is reminded of a similar model for defining *minimal chemical life* as consisting of a programme, metabolism, and a compartment (or "PMC").[61]

What I find interesting with this contribution is the admission that they are descriptions and not explanations. They are claiming novelty, but in property and system-level description. They know that if they claim some other kind of stuff they would need to, for example, qualify it with a range of micro-level and environmental conditions over time. However, they still fail to recognise the metaphysical significance of the system's dependence on the environment (considering the feedback and throughput mechanisms they admit to), as well as the notion that the system qua whole must exist over time in order to have the system effects it describes. The system-whole cannot be static and synchronic as these and the previous (not to mention other) authors are assuming.

[61] This is based on Bedau's recent work on proto-cells (2009), which was inspired by Maturana and Varela (1988).

1.2.7. Autopoeisis and Autocatalysis [62] [63]

Throughout their writings, Maturana and Varela (1980, 1988) are explicit about their use of biological systems qua *unities*. The authors suggest that, because cell reproduction involves self-replication and the generation of another cell of the same class, the replicator must be unified in order to replicate. This is a solid argument in some ways, but not in others. The phrase 'of the same class' is the source of the problem: it is neither numerically nor qualitatively identical, yet shares many of the same properties. So, the cell qua unity simply finds a way to replicate the very processes that manifested itself in a similar way as conditions allowed. This is the essence of autopoeisis, where a living thing is understood as self-creating and self-maintaining. The obvious question is: how unified can something be, in the metaphysical sense, if it is condition-dependent, and replicating, copying and reproduction are never perfect? (1988, 59). Nevertheless, the authors would consider the cell as an autonomous unity, and its self-creation reinforces the idea that the cell exists qua whole.

Autocatalysis in biological systems is when the product of a chemical reaction is also the catalyst for that reaction. In various books and articles, Stuart Kauffman has nominated autocatalysis as a candidate for both abiogenesis (how biological forms came from inorganic matter) and the origins of what he calls 'agency'.[64] He uses self-organisation to show that life (strongly) emerges out of concentric circles and vast matrices of autocatalytic reactions, and hints in *Investigations* (2000) that a broader synthesis of autocatalysis and self-organisation in what he calls '*coconstruction'* may

---

[62] In the context of abiogenesis, Kauffman is not the only proponent, as Richard Dawkins has given the notion some treatment and support in *The Ancestor's Tale* (2004). Also, Kauffman's major contributions on this and their relation to emergence can be found in *At Home in the Universe* (1995) and *The Origins of Order* (1993).

[63] Related to this topic is the work of Protevi (e.g. 2006, 2009), who uses Deleuze and DeLanda (e.g. 1997, 2002) to conceptualise something called "transversal emergence": the production of distributed and interwoven complex systems or assemblages through iterations of interactions. Basically this is an argument for the *de facto* ontological emergence of really big, messy, dynamic things that are (typically) organic or social. For me this material should be presently sidestepped due to its (a) problem of individuation as a candidate for a new category of existence, and (b) deeply anthropocentric lens. If it cannot be individuated, then no properties can be ascribed to it, let alone be considered ontologically novel. If it were to be a case of emergence *simpliciter*, where property ascription is possible in light of its instrumental individuation, as we will see later in this thesis, it still is ineligible for ontological emergence.

[64] This is one of the topics in *Investigations* (2000).

help us to understand these ideas in a cosmic setting. The collective work is the most prominent complexity approach amidst the biological sciences, and is justifiably amongst its most important contributors overall. Though we cannot possibly give a full treatment of Kauffman's work here—and I do find the co-constructive notion particularly attractive—I will briefly discuss issues surrounding his main unit of analysis: the self-organising system (SOS).

First, the term 'autocatalysis' reveals an important element. How can something be both the catalyst and product of the same reaction? There are two ways to go from here. (1) In order for something to be both, it would need to take place over time—otherwise it would be impossible to distinguish between the two observationally. If the events that led to a reaction are dependent on the result of the same reaction, the reaction itself must have a sufficient duration to permit circularity. The time that it takes for the tip of the reaction output to feed back into the base of the reaction input must be less than the entire duration of the reaction. In most cases, it is a cyclical and indefinite process. This is only possible if the entire system is dynamic and diachronic, as well as exposed to suitable environmental conditions.[65] (2) Secondly, this confusion between different start and end times is quite possibly caused by a confusion of agency. What do we mean by reaction product? Perhaps its constituents are not entirely fixed, meaning that agency is not fixed, and the ascription of cause and effect is misleading. It might lead us to look at SOSs as being a misappropriation of identity, in philosophical terms. When someone claims that a SOS is a singular or collective autocatalytic set they might be simply calling it that for literary convenience—or maybe for parsimonious, experimental reasons. We might consider to have in Kauffman (at least in the more technical writings, e.g. *Origins of Order*) a limited set of metaphysically-significant claims, and it might be shown to be unhelpful or even impossible to treat many of his claims in the metaphysics of emergence. Nonetheless, the implications of at least some of them are metaphysically significant.

According to Kauffman, the SOS is an *isolated* system: this is to say that one may be able to claim that no matter is exchanged with the environment, but one cannot have

[65] I have tried to downplay the role of ontological dependence throughout this chapter, so that I may introduce it gradually later. It is difficult though to avoid doing so in certain cases along the way.

a full account of the force fields and particles coming and going. It is certain that a biological SOS exhibiting autocatalytic features could not exist in a closed system: this is simply because in an experimentally-closed system we have only ever found the 'laws' of thermodynamics to be perfectly true when testing on a specified level of analysis.[66] Most importantly in an experimentally-closed system, nothing can be created or destroyed (#1) and disequilibrium always increases (#2): this means that something must account for the spontaneous self-organisation of a system, as well as the spreading out into new, enlarging "workspaces" (2000). If the workspace expands, which is a necessary condition for large-scale autocatalytic sets, the system can no longer be considered an isolated (let alone closed) system. Furthermore, in order to co-construct and co-create, aggregates of matter and force must be fused with other such aggregates in order to achieve either 'equilibrium' or the balance between chaos and order that Kauffman likes to ascribe to SOSs. And this is all taking place, by Kauffman's own admission, in the context of a singular unit of analysis. Yet, if the SOS is compositionally and operationally dynamic and diachronic, how can the unit of analysis remain unaltered? How can we ascribe intrinsic properties to a thing that has by definition no fixed state? Individuation becomes a parsimonious exercise, and not an empirical one, or properties at certain microphysical levels would need to be considered irrelevant to their respectively higher levels.

In plainer terms, I believe we should consider the main thrust of Kauffman's co-construction and co-creation as a wonderful philosophical tool for describing in macro-level terms a broad view of the cosmos (or at least biospheres). And, perhaps he is not interested in any sort of intrinsic properties, nor is he attempting to contribute to the discipline of metaphysics. But, he clearly favours the basic notions of holism and strong emergence, stemming from his idea that there are causal features of biological wholes that their microphysical constituents cannot have.[67] I think he has misplaced his trust in the role of agency and macrocausation that inevitably downplay the role of

[66] This is to say, thermodynamical laws apply to isolated or experimentally-closed systems: they are not really closed, but they only test for certain things, not requiring, for example, measurements of sub-nuclear reactions. The costs of measuring and including such micro-variables would be too high for the myriad of tests that go on to reinforce these thermodynamic principles.

[67] The (strongly) emergentist leanings are more obvious in *Origins of Order* (1993) and *Reinventing the Sacred* (2006).

the microphysical[68]—something that self-organisation, in the term's own meaning, proudly declares. In order to claim the *true autonomy* of a system that purportedly organises itself, achieves equilibrium, and strikes a balance between order and chaos, it would need to assume the existence of enough food and energy, a high enough concentration of its own molecules, and a fixed probability for chance catalysis. At a minimum, these chance conditions—rarely offered by the cosmos—must be met in order for the mixture of molecules to be able to transform itself into a stable and self-sustaining set.

Kauffman's emergentism would generally be considered strong, being more akin to Van Gulick's taxonomical description (and not claim) of a 'radical' kind—after assuming features included in the 'modest' kind. I hesitate to say that he wants to claim a new kind of stuff apart from physicalist bits, but he clearly is in the downward or at least macro-level causation camp. Nevertheless, the arguments, when roughly tendered in metaphysics, suffer from similar gaps as above, as he is a prime example of an emergentist whose crucial flaw is shared by his reductionist counterparts: the unit of analysis is operationalized as a whole, but its mereology is never elucidated. It could very well be that it never really was a whole or that its integration was never complete—let alone ontologically independent—so we should not rush to treat it as autonomous, essential, and/or holding intrinsic properties.[69]

[68] Prioritising the 'microphysical' does not necessarily signify the priority of physics or any of its sub-disciplines. All it necessarily signifies is that our arrow of explanation is presumed to have, in geometrical terms, a negative slope (looking downwards).

[69] In Clayton and Kauffman (2006), the authors—despite arguing for molecular autonomous agency—admit a) that there is no "general theory of the organization of biological processes" (thereby allowing some room for emergentist explanations) and that b) biological agency is to be explained in biological terms (indicating level-exclusivity). Interestingly, in the face of such explanatory uncertainty, they also seem to indicate that, agreeably, biological emergence (autonomously molecular, irreducibly weak and epistemic, or otherwise) is a necessary part of investigation. Unfortunately, they do their general argumentation a disservice when they try to claim that biological species extinction is evidence of downward causation, making the false assumption that a genome is a static entity.

1.2.8. Dupré's anti-reductionism

John Dupré and various Stanford School members[70] have long been critical of reductionism in most if not all of its forms, though their proposed alternatives might not really be considered as emergentist *per se*. I do, however, want to take a brief look at an attempt on his part to challenge what we are now calling ontological reduction.[71] The author is what we might call an ontological pluralist (e.g. 1995): there are many unique things in the world, with their own causal powers, with no possibility of property, causal, or ontological reduction. It might be described as a kind of holism, with mountains, markets, plants, animals, and humans being among those things irreducible in any way. In a general sense, their system-level or higher-level properties are emergent because they are not resultant from the properties of their parts. There is less an emphasis here on the *transition* to novelty than there is on plain difference, and this is why it is often not associated with emergence.

In looking at one particular excerpt, Dupré has taken the liberty of ascribing system level and even causally-efficacious properties to a *street market*, but not to an automobile. His following description of the car in reductionist terms is meant to illustrate an insurmountable gap between it and an anti-reductionist description of wholes like the ones I have already mentioned.

> The parts of a machine are not unchanging, of course, but their changes constitute a relentless and one directional trend towards failure. A good machine starts with all its parts precisely constructed to interact together in the way that will generate its intended functions. The technical manual for my car specifies exactly the ideal state of every single component. As friction, corrosion, and so on gradually transform these components from their ideal forms, the functioning of the car deteriorates. For a while these failing components can be replaced with replicas, close to the ideal types specified in the manual, but eventually too many parts will have deviated too far from this ideal, and the car will be abandoned, crushed, and recycled. […] Reductionism is almost precisely true of a car. We know exactly what its constituents are – they are listed in the manual – and we know how they interact: we designed them to interact that way. (2008, 13)

[70] Notable works include Nancy Cartwright's *The Dappled World* (1999) and *How the Laws of Physics Lie* (1983) and Ian Hacking's *Representing and Intervening* (1983).

[71] He does not separate the varieties of reductionism categorically; rather, he tends to use a broad all-encompassing label to refer to methodological, epistemological and ontological 'forms' of it.

Contained within this quote, I believe, are some of the most common sources of tension and possible misconceptions of reductionism in the physical sciences, some of which have already been mentioned here in §2. There are three points: one technical, referring to this specific passage, one mereological (or compositional, or ontological), and one on properties (potentially epistemological).

First, there is no thermodynamic rule that says any machine will *necessarily* decline: nothing exists in reality as a closed system, as interpreters of the Second Law often forget. So, the 'parts' could in fact be changed indefinitely—though we would end up with the identity problem of *The Ship of Theseus* (to be discussed in Chapter IV). The directionality of decay and decline can only be cosmic—meaning beyond any imaginable scale—and linear, reversible time belongs to Newton, not contemporary physics. All of this means that we should not expect the parts of the car to decline unless we have reason to believe we will know the conditions within which the parts will exist. On this planet, in a temperate climate, for example, we can expect certain levels of temperature, pressure, and precipitation, amongst other things, to affect the components in an estimated way. We frequently calculate the number of miles that a given part is expected to perform for, assuming normal usage and conditions. If conditions were more severely monitored and (imperfectly) controlled, we could easily extend the life of any number of given parts. And, this bears a striking resemblance to maintaining the health of an organism.[72]

Second, the parts of the car that we design and install are not the only parts of the 'system'. Amongst many other things, the car relies on suitable fuel and air to maintain smooth combustion, as well as a range of suitable environmental conditions in which to run. Variations in these conditions will obviously affect the performance of the parts as they have been designed and installed. Hence, the composition of the car has been hastily ascribed, and the consideration of the car as a whole (or at least as a discrete object) without autonomous causal powers (because they, like all machines, can be reduced) reveals a bold selectivity about which parts and properties to include in

[72] I will return to these issues in Chapter III, but one thing to keep in mind from §2.7 and §2.8 is how evidently common it is to grant special status to conditions and levels of analysis closer to our own individual human level—things that we are most familiar with.

descriptions and explanations. When we look at this claim in a different way, as revised according to §1.0, we see the following:

- (1) *Revised*: Car C = $\{o_1 \ldots o_n\}$ = Car C is composed exclusively of parts $\{o_1 \ldots o_n\}$.

But, Car C, at any given observation, does not consist exclusively of the parts listed in the manual, or $\{o_1 \ldots o_n\}$. Car C, even instantiated at t_1, cannot be synonymous with $\{o_1 \ldots o_n\}$, because there are other parts to the system not listed in the manual $\{o_1 \ldots o_n\}$. Because, for example, oxygen is a necessary component for combustion, it must be included in the set $\{o_1 \ldots o_n\}$, but since it is not in the manual, the set $\{o_1 \ldots o_n\}$ is incomplete. And, this only begins to scratch the surface of this problem of hasty integration.[73] In fact, the system C, *especially* from the perspective of Car C qua causally-efficacious autonomous agent, is only *realizable* when all these other microphysical parts $\{m_1 \ldots m_2\}$ come together under the right conditions. Therefore Car C is not synonymous with $\{o_1 \ldots o_n\}$. Rather it could be stated as:

- Car C = $[\{o_1 \ldots o_n\} + \{m_1 \ldots m_n\}]_Y$ [74]

It seems to me that this, as simple as it is, negates the claim in (1) *Revised*. Now, if we were to anticipate a rebuttal, it might come in the following form: "Why should we consider gases, heat, and corroded particles as parts of the whole, when they do not linger long enough, nor are they recognisable enough, relative to their identity within the system?" This is what I am calling 'level exclusivity', and it is what enables pragmatic explanations (and applications) in engineering. It cannot, however, contribute to any explanation in the metaphysics of emergence. To question the mereological horizons upon which our wholes and units of analysis have been

[73] The car's fuel, as well as the steel parts themselves, must also be able to exist within a range of conditions. To take one example, conventional steel (an iron alloy) melts around 1370°C, so if Car C were subjected to such conditions, $\{o_1 \ldots o_n\}$ would not have the same properties, and potentially threaten their individuation as parts of C—meaning, C could not be integrated and, eventually, individuated. C would cease to exist, or at least would lack observational evidence for being an individual object.

[74] And, actually, this could be even further restated to reflect the condition-dependence of the observer at a given instantiation: $\{\text{Car C}\}_{Y1} = \{(o_1 \ldots o_n) + (m_1 \ldots m_n)\}_{Y2}$. The transcription of this would read: observation of Car C under (imperfectly-detected) set of conditions Y_1 show us that Car C at t_1 is composed of proper parts $\{o_1 \ldots o_n\}$ in combination with microphysical components $\{m_1 \ldots m_n\}$ all existing under the (imperfectly-detected) set of conditions Y_2. Though awkward, this formulation is, to the best of my knowledge, precisely how a physicist would see the problem—and I believe this is where we should be *starting from* in the metaphysics of emergence.

constructed is the very purpose of this thesis, and I hope to show gradually that, whilst we *can* exclude the extreme microphysical in parsimonious descriptions, we *cannot* exclude it in metaphysical explanations. This of course means that, in order for them to be valid, claims for individuated novelty (such as claims for strong emergence) must acknowledge themselves to be empirically-sound claims for entities that are both unified and persistent.

There is a final point here on properties. Dupré and many others like him will, with good intentions I think,[75] go on to say that, unlike machines, life forms (and/or their properties) cannot be reduced to (the properties of) their parts, where unique system-level properties combined with autonomous causal capacity prevent us from explaining the forms according to a "manual." But here is another problem with this idea: an autonomous, irreducible, emergent, system-level property must also be an *intrinsic* property. First, a definition: this is a property ascribed to an object irrespective of any other object, or causally independent of any external physical aggregate. Now, such a property is conceivable in a worldview that includes forms, ideas and essences as irreducible entities, joining the ranks of matter and energy as fundamental categories of existence. But in an empirically-sound framework, where physicalist bits are ontologically primary, this does not seem to stand up.

As to why such a system-level property must be intrinsic, I believe we should consider the notion that an intrinsic property first presupposes unity, individuation, and probably category (and, as we will see later, individual objects' persistence). If property *P* is intrinsic to object *O*, or to a class of objects *OO*, then *O* holds the property *P*. If *O* holds the property *P*, so must the constituents of *O*, or set $\{o_1 \ldots o_n\}$ also hold the property *P*. Individual members of the set need not have the property P, but the set, as when taken together, must. Ascribing (or identifying, for that matter, depending on the worldview) an intrinsic property to an object *demands* mereological integration of some kind, for

[75] I say 'good intentions' because the basic ethos of the post-structural anti-reductionism is predictably responding to the arrogance of the physics-driven quasi-political agenda of (pre- and) post-war science practice—often being associated with the abuse of science and questions of ethics. Obviously this is far outside the scope of this work, though I mean to express my own sympathies with attempts to preserve a certain level of dignity for the organism in the philosophy of science—particularly in the second half of the 20th century.

when we say *O*, we mean "whatever *O* is and whatever *O* is made of." Hence property P is applied to O as an individuated and unified object.

To view the car as a metaphysically unorganised heap of elements, possessing exclusively extrinsic (and not intrinsic) properties, and without unity, is not *unacceptable*. It is, in my view, rather acceptable. What does not seem to carry much weight in an empirically-sound physicalist worldview, is the view that life forms, for example, are different because they can be seen as individuated unities with demonstrable, supersummative agency. This has long been the goal of emergentism: to see organisms as seemingly occurring 'naturally' or spontaneously, and as vessels for intrinsic properties, or possessing an innate proclivity to agency—e.g. downward causation. Should the system-level and self-organising properties exhibited by, for example, mammals, be synonymous with intrinsic properties? Are they properties existing without external reference, where they merely *are*? With these questions, I think it becomes clear that strong emergence inexorably connects to properties, and so it must also connect to epistemology.

1.2.9. Strong emergence reimagined

With most of the authors above, weak claims are the stepping stone to strong ones, and it is the strong claims that are of interest here. What is a strong claim? Is it when there is a new category of stuff? Is it when there are new, irreducible causal powers, or when there are new, irreducible downwardly causal (DC) powers? Is it perhaps both? Does DC entail a new kind of stuff? Silberstein and McGeever seem to conflate some of these questions, and define ontological emergence as:

> [When] features of systems or wholes [that] possess causal capacities not reducible to any of the intrinsic causal capacities of the parts nor to any of the (reducible) relations between the parts. Emergent properties are properties of a system taken as a whole which exert a causal influence on the parts of the system consistent with, but distinct from, the causal capacities of the parts themselves.
> (1999, 182)

This statement muddles weak and strong elements, just like the progression from Kim and Ellis, where later stages presume and incorporate the earlier claims. 'Ontological' seems to have taken on wildly different meanings going beyond 'category or class of existence'. It seems to me that there should be a distinction between *ontological* emergence and *strong* emergence, because the word 'ontological' should, I think, be reserved for categories of existence. If something is ontologically emergent, it must signify a new kind of fundamental stuff apart from that recognised by physics. Physics recognises matter and force in a standard model—each broken down into different manifestations—so if the stuff is not made of matter and/or force in any of these manifestations, it must be a new kind of stuff. This claim is evidently much harder to make.

I believe it is a mistake to conflate evidence for autonomous causal capacity with evidence for ontological novelty. The ability of an object to create change, to be an autonomous agent, or, more precisely, for it to instantiate new properties in the world, should not necessarily be associated with a new category of existence—or an object composed of things apart from matter and force aggregates. If 'causally efficacious' becomes interchangeable with 'ontological', then the entire physicalist foundation becomes threatened—and potentially meaningless. The ontological domain must be reserved for fundamental existence: *evidence for* autonomous causal powers can be used to suggest an addition to the physicalist base (and this must be considered as a distinct possibility), but observation and coding of such powers does not make the issue immediately ontological. In other words, if a given object exhibits properties for macro-level agency, we discuss those properties within a physicalist system knowing the limits of common human observation, sensory data transmission, coding, description, and attempts at explanation. We do not instantaneously cross over into ontology because we ascribe some properties of self-organisation and, potentially, autonomous agency.

What is perhaps more important on this point, though, is the fact that new causal powers *do not actually necessitate* a new category of existence. The ability to instantiate new properties in the world does not require object *O* to be something *more than* matter and force aggregates. This is ultimately because a challenge to the standard model is much more significant than a challenge to common, unaided, and untested observation (and its coding, etc.), and I assume this point is uncontroversial. Analysis of the objects in question may indeed call for such challenges, and ontological emergence is, again, a distinct possibility, but the road to verifying the existence of supersummative agents is empirically a long and difficult one. I believe the only way to enter into ontological discussions in the metaphysics of emergence is, initially at least, through direct claims for new categories of existence—which again, are not, in principle, off the table in the metaphysics of emergence.

Strong emergence should, I think, refer to causal powers of the whole which do not appear to derive from that which is contained within the whole. This means that the system exhibits unique, unpredictable causal powers that individual parts do not have. The recipient or passive agent of these causal powers might either be a) another object (at least one), where the whole exhibits macrocausal powers (MC). The recipient can also be b) other objects as well as its own parts (DC). On lower levels of physical reality, the causal powers of a unit of analysis typically seem to be derived from its properties. Properties, however, must be ascribed to objects, which can then be considered in some cases to be wholes. Several authors seem to think that DC either signifies a new kind of stuff (rare, mostly in pluralist ontologies or mind-body dualism) or at least must be included in *ontological* discussions. If the claim does not involve a new kind of stuff, I would argue that DC should not be included in ontological discussions.

The second claim b) is where strong emergence claims become equated with claims for DC. MC is taken for granted in most weak and all strong claims, as well as the claims for ontological emergence as I have just re-defined it (and it is this assumption that I will return to in the next section). Strong claims must first claim MC, just as in most weak claims, but then they must also go further and claim DC. The controversy of this reworked strong emergence concerns DC: the potential ability of the whole to control its own parts. The prime example forwarded for this is the mind, where higher brain functions on a system-level, such as mental events, seem to be able to re-arrange the synaptic patterns and biochemistry that in principle must comprise the same system. I will look at this MC/DC problem in Chapter III.

None of the key authors actually directly and overtly support or claim ontological emergence in the sense that they have unearthed a new kind of stuff. Even Silberstein and McGeever, who go on to argue that non-locality in quantum mechanics is ‘strongly’ emergent, end up taking the DC pathway. The only authors espousing such a position are those found advancing a substance dualist or pluralist ontology, and they are not the primary target of this work—though perhaps still sharing the same misapprehensions of the ‘whole’. In the models above, the authors generally address a progression or gradation between levels of complexity and causal powers *on the system level*, so they are in the strong (and not ‘ontological’) camp to which I have, for the sake of clarity, reassigned meanings.

1.3. Mereology and diachronic emergence

What has been assumed by many in the general discourse of emergence and most of its applications throughout the humanities and social sciences (though not in the metaphysics of emergence) is that strongly emergent objects and properties cannot be described or explained using (previously) standard models of time, persistence, and physical (de-)composition. More specifically, ontological and even property reduction are not only untenable *after observation* of these candidate emergent objects, but also through the nonlinearity and non-uniformity of the object’s and properties’ temporal existence. Strong or ontological emergence is a real feature of reality, according to this assumption, also because, where it exists, there are multiple levels of reality whose constituents cannot be individuated because of an inherent *horizontality* or horizontal dependence. The candidate emergent object (either in stuff or in having such a property) is dependent on multiple tiers of lower-level objects, events, or properties that exist dynamically, that make the macro-object or system equally dynamic, and, as such, that render it ineligible for instantiated analysis. The higher level, or the system-level, is the only means to capture this dynamism, and this exclusivity can generate, amongst other possibilities, arguments or evidence for irreducible causal powers.

This assumption seems to be a challenge for not only the ontological and reductionist position, but even for the reintroduction of mereology into strong emergence. Why? A mereological dimension to strong emergence would assume and require at least sporadic compositional instantiations to understand the relationship between parts and whole within the analysis of a candidate emergent object. These instantiations would necessarily require us to freeze an object in order to see where powers and properties come from. The original understanding of diachronic emergence (as described here) assumes that these lower-level tiers are shifting and bubbling through time at different rates, with each moment blurred by the next—thereby denying the ontological use of the instantiation. If snapshots of emergent objects are impossible, then how could we even begin to talk about emergent wholes, which must have a finite and discrete composition at at least $time_t$?

First of all, this is actually a topic that is addressed throughout the thesis, and because of its centrality, there can be no tidy resolution to this in Chapter I. However, there is an obvious digression that will not be critical to the work in its own right that should briefly be included here. Fortunately, there is enough overlap with the core material such that this next section could prove to be useful.

1.3.1. Process and event in strong emergence

The dynamism of many candidate emergent objects is incontrovertible, either compositionally, ontologically, observationally, or especially with regards to properties and powers. The diversity of their constituent objects and properties (and their conditions) possessing this dynamism incorporates countless levels of analysis, including those beyond the scope of unaided observation. But, what about constituent objects and properties that do work, but are ultimately accelerated by same- or lower-level *events*? Earlier we said that an event was the interaction of two or more objects, necessarily with a duration greater than zero, that is itself individuated through properties, that can itself generate event-level properties, and where the interactive objects must themselves be discrete objects (though not necessarily static). Would it not be tempting to make a special place for *event* in any discussion of strong emergence? What is the relationship between *event* and *process*, assuming the latter holds significance beyond our description of *event* used above? Does it make sense to say that an *event* can be a *whole*?

These questions are essentially answered throughout the thesis. For now, I will set the stage for doing so—particularly in consideration of what I have already discussed in this present chapter. First, our description of *event* suggests that it cannot be ontologically primary, because it is merely a description of interactive object behaviour under certain conditions with a finite duration. So, to know an event is to know the constituent objects and their conditions—which need not occupy the same level of reality. Any examination of candidate emergent objects is focused on the question of whether the object is or has properties or powers that are unpredictable and/or irreducible (et al). The question is *not* whether an event which permits the instantiation of a constituent is emergent, or whether the candidate object is to be considered as emergent (in some way) *because* of this event. The question is whether the object and/or its properties are emergent.

So, whilst events (which may horizontally occupy space and time before and after the instantiation of the candidate object) may enable or compose the candidate object, the metaphysics of emergence (at least in our present investigation) cannot be concerned with constituent events as either ontologically primary or eligible for emergent property ascription. Strong emergence must—eventually—turn and face the question of the object, even if the object has all kinds of dynamic things happening inside and through it. And, as it turns out, as we progress through this work, we will see that the dynamism of and enabled by various currencies (especially the event) proves to be critical to assessing the validity of strongly emergent claims. But, it does not start with the event. It starts with the object. It starts with the (potential) whole.

So is there any meaningful reason to distinguish between *event* and *process*? Aside from an obvious digression on the history of the use of the word and its corresponding process philosophies and schools of thought, I do not believe much is to be gained from distinguishing the two here—with one brief exception. An *event* is often understood as possibly being unanticipated, unpredictable, and/or spontaneous. These adjectives all have different specific connotations, of course, but they all seem to contrast with *process*. A process seems to be some kind of event that has been observed before and, conceivably, attempts have been made to understand more clearly what happens when it takes place. In other words, an event can simply happen, but a process seems to require the accompaniment of an observer's insight. This is why I prefer *event*: it shoulders no pretence of how precise we think our understanding of an event is. Of course, for some of these process schools, the very word *process* is an ontological primitive, so it carries some special kind of reverence. Nevertheless, in this work, I use it essentially interchangeably with *event* and reapply its description.

Also, on the question of whether an *event* can be a whole, it seems to me that wholes can only be objects—though we will see more of why this is the case in Chapter II. An event's spatial and temporal limits are restricted by the observer, but the hunk of matter-and-force particles/fields that we will come to call an object, is not. For example, (the structure that we call) a hydrogen atom's existence is not dependent on our observation of it. But, its (presumed eventual) bonding to a nearby oxygen atom (as permitted by local microphysical conditions) requires an interaction between the two, and that interaction has, necessarily, a duration greater than zero (by standard gravitational conditions). The precise duration of this event and the rationale for determining t_1 and t_2 is entirely dependent on an observer's whim. This relegates an event to, essentially, 'what we create when we want to track multiple objects.' Furthermore, even though the observer can decide what an event might be, the observer has no say in what could happen 'after' this forged event has transpired. Does the new molecule add a second atom of hydrogen, and produce water, forming the substance that permits all known life? Or, does the new molecule form new covalent bonds to form alcohol, which can be deadly to countless cells and microbes? How is the observer to know whether their originally-constructed event was a mere fragment of a much larger event? Or, could this selected event be an aggregation of hundreds of other events? Again, this generates an endless list of questions leading to metaphysical vagueness, despite the importance of *event* in observation, description, and explanation. But, the observer is off the hook for these questions, because focusing on $event_E$ precludes them. The ascription of the property irreducible (or unpredictable) is possible only with reference to objects, because strong reduction of events must first pass through its constituent objects. The metaphysics of emergence, as I see it, then, has to focus at least initially on the object and its potential wholeness and irreducibility.

1.3.2. Process philosophy and strong emergence

Inasmuch as *process* opposes itself to *substance*, this thesis could very well see itself as having a process-based lean—but only through our 'refined physicalist' lens. That is to say: not much, if anything, in the natural sciences today, assumes static composition. Even gene-centred philosophy of biology now often prefers to have an epigenetic lens—where gene expression is dependent on local microphysical conditions. Process philosophy, in its supremely generalised form, therefore, has elements which are entirely uncontroversial especially in the domain of the metaphysics of emergence.

Substance metaphysics is really just classical metaphysical currency. It is the sterling coin to the algorithmic stock purchase. Without a fair grounding in what real-world objects (as well as events, properties and conditions, constituent or otherwise) are doing, dynamism seems to be bereft of meaning and incapable of philosophical traction. It is quite easy to dismiss that which seems to have no connection to our own evidently-biased view of reality (for example, there are contemporary metaphysicians who seem to assume that atoms are intrinsically stable entities). But, it is no longer a matter for debate to say that things *emerge* (in some way), and that no physical object is immune to at least microphysical change. In other, more diplomatic words, substance metaphysics can no longer be an exclusively-empirical domain. Therefore, its classical rival *process* seems to have a fairly important seat at the table.

And yet, unlike *substance*, it is far too disorganized to even justify applying a singular label. There is classical process, such as Heraclitus or threads of ancient atomism (discussed in Chapter II), and shards of process scattered ever since (like Plotinus and Leibniz, also Chapter II)—often overshadowed by the dominant substance paradigm. Of course Whitehead is now the primary face of process philosophy, though his contributions to contemporary (refined) physicalist metaphysics (and this thesis) are limited. So, which contributions of this central figure are noteworthy for our purposes (noting that a full treatment of Whitehead's philosophy falls outside the scope of this work)?

1.3.2.1. Whitehead's resonance

Rather than substance, the event, or "actual occasion",[76] is the singular ontological primitive. Rather than looking for objects in the world, the focus is on the event, which ostensibly manifests the things we call objects.

> An object is an ingredient in the character of some event. In fact the character of an event is nothing but the objects which are ingredient in it and the ways in which those objects make their ingression into the event. *Thus the theory of objects is the theory of the comparison of events.* Events are only comparable because they body forth permanences. We are comparing objects in events whenever we can say, 'There it is again.' Objects are the elements in nature which can 'be again.' (1920, 143-4, emphasis added)

If objects are at the mercy of enabling events, then it is not clear how respectively lower-level objects constituent in the enabling events could ever hope to compose an event. If events are the sole source of agency, then object individuation would be irrelevant. If object individuation is irrelevant, then how are we to distinguish between two events? In other words, events can be known by their constituent objects, but their constituent objects cannot be known without recourse to their encompassing event. This seems to be a problem, though it is really just a reformulation of the same problem found in substance metaphysics. Nevertheless, what I believe is useful here, is not only the acknowledgment of the inherent and insurmountable dynamism of lower-level processes, but the porosity that Whitehead assumes. Even if his objects are (unacceptably) reproducible through irreproducible events, the events that generate these seemingly static objects are, for him, necessarily nominalistic. This leads me to believe that Whiteheadian events are actually just what I have been calling *conditions*: topographical approximations of object behaviour and related same- and lower-level events.

[76] An actual occasion is the integration or "concrescence" of data transfers (or "prehensions"), such that each event is the aggregation of cosmic information and ultimately a process of becoming (1929). The only salvageable notion that I can determine from the specifics of the concept of the actual occasion is its resonance with the laws of thermodynamics (pertaining to isolated systems), which would have been slowly becoming solidified in his time. Additionally, much of his process metaphysics was actually focused on how to see some kind of god-entity as a process in and of itself. This is obviously not a focal point here, or within a refined physicalism and empiricism. My interests lie exclusively in sketching an understanding of *event*, using various approaches.

Individuating events is the issue for me. This is not to say that individuating objects (which is perhaps the key dilemma of this thesis) is uncontroversial. But, individuating objects has fewer barriers, and requires fewer assumptions, than individuating events. In order to have a theory of objects, we must, as he says, compare events. Now, in any investigation involving multiple levels of analysis, there is always the threat that higher-level description and explanation is dependent on lower-level objects, events, properties, and conditions. Making the *event* ontologically primary means that every respective constituent event is increasingly less important than its constituent event: in other words, if the lines between individuated objects were blurred by their enabling event, then the lines between enabling events are even more blurred. It is not merely a problem of infinite regression (which we will discuss in the next Chapter) for object (physical and causal) composition, but an even more troubling problem of infinite event regression. Because events have durations greater than zero, knowing their bounds is a matter of temporal precision—which we can never verify. Ultimately, the event's dependence on lower-level events is, at first glance, the same problem as ontological reduction. However, it turns out that this dependence is even sloppier because it requires temporal precision (11:59:00:00:08 → 11:59:01). An ontologically-primary event may have extraordinary explanatory power, but it is always a level-specific explanation. The object, as I hope to show, is not limited in this way.

I might go as far as to argue that Whitehead is more Aristotelian than he might think. We have already mentioned the shared problem of individuation in (at least) the microphysical domain, as well as the level exclusivity that plagues much of the history of metaphysics. To those problems I would add what I see as the great classical blunder of seeking meaning and purpose in that which we now know to be purely matter and force particles and fields interacting with local conditions—without perspective or function.

> There persists a fixed scientific cosmology which presupposes the ultimate fact of an irreducible brute matter, or material, spread through space in a flux of configurations. In itself such a material is senseless, valueless, purposeless. It just does what it does do, following a fixed routine imposed by external relations which do not spring from the nature of its being. It is this assumption that I call 'scientific materialism.' Also it is an assumption which I shall challenge as being entirely unsuited to the scientific situation at which we have now arrived. (1925, 22)

Of course, when he speaks of this "irreducible" matter, he is speaking in reference to the physics of his day, when the present standard model was non-existent. His day was, in fact, just beginning to digest the validation of the concept of the atom as a real class of entity through Einstein's Brownian motion. Nevertheless, he clearly *wishes for* an organic view of nature, and he tries to make it so by constructing an elaborate scheme of organized dynamism that, in some writings, is unapologetically teleological. He seeks to impose a structure of processes so that he can make sense of purposeless matter. We cannot fault him in that general quest, but unfortunately, there is no evidence to suggest existence is little else.

1.3.3. Diachronic lessons

As I have said at the beginning of §1.3, mereological approaches to strong emergence are seemingly challenged by an increased importance of *event*. But, let us be clear about what it is we are trying to do here in this work: because the validation of claims made in the metaphysics of emergence is dependent on using discrete units of analysis, and since those units must necessarily be individuated based on assumptions of integration, the metaphysics of emergence is necessarily a mereological—and therefore object-based—exercise. It requires integrated discrete wholes (which are objects), and is composed of other objects, events, properties and conditions. For those constituent currencies, they must be discrete, and they must be confined to occupying the same region of space-time as the whole in question. If we are required to use instantiations, the whole's individuation cannot be sacrificed.

And, as it may turn out, it could become clear that either this approach is untenable—which would support the diachronic challenge—or that there is another way to measure and assess the claims made in the metaphysics of emergence. In any case, one thing is certain: dynamic and diachronic emergence is an empirical matter, and incorporating it into a system of thought is not a matter of whim. I simply suggest here in Chapter I that there may be multiple ways of understanding it (so-called diachronic emergence), and therefore using it as a tool for description and explanation.[77]

The importance of the event is staggering; and yet, as I see it, it is subordinate to the object. This is because, again, events are composed of objects, and objects, as hunks of matter and force particles and fields to which properties can be ascribed, are the only currencies not *requiring* an observer bias. Properties require our encoding. Events require us to specify durations as well as agency ascription. Conditions require us to determine spatiotemporal boundaries as well as which objects and events are relevant.

Furthermore, it is important also to see how something can be both discrete—like a whole—and not static. First, there is level-exclusivity, where we can say that the effervescence of an object's microphysical constituents is irrelevant to the higher levels (this will be addressed in Chapters II and III). Second, we can say that a whole has an essence or founding relation to not only bind the ingredients (or parts) but also that determines which ingredients can be safely lost (Chapter II). Thirdly, using wholes in the way that we will depends heavily on the use of scattered instantiations, where the analysis of which cannot include anything other than momentary snapshots. Challenging their usefulness is putting us back at square one. Using instantiations has no bearing on the object's static or dynamic behaviour: it is merely a strategy to freeze time and see where things line up. Using wholes allows us to a) treat claims in the metaphysics of strong emergence as they must necessarily be treated, and b) incorporate the dynamism that any constituent object, event, property, or condition might present us with.

1.4. The necessary mereology in the metaphysics of emergence

A central aim of this chapter has been to look at some of the primary claims for strong emergence and question some of the uses of 'object' and 'whole' in order to emphasise

[77] Again, this is central to the whole thesis. It should be obvious that a so-called 'diachronic' emergence is really just 'emergence', and I fully treat this throughout. But, in particular, a close reading of Chapters III and IV should eliminate all doubt.

the profoundly important role such a definition has. Unfortunately, we almost never get a definition of a system, an object, or a whole, in metaphysical terms. There are references to wholes, but hidden assumptions always seem to linger about their actual nature. I will now clarify why the objects in question—the objects to which these authors are considering strongly emergent or reductionist property ascription—all seem to converge on a singular totality and whole.

1.4.1. From the authors

Searle weaves into his writings references to entities, objects, wholes and levels. The phrase "underlying levels" is used repeatedly to indicate the existence of both a higher and a lower level (presumably of complexity). He does not make an attempt to distinguish between them and seems to use them interchangeably. In this case, a higher level might be labelled "highest level", as it encompasses all levels as the entity, the object, and the whole in question.

Van Gulick restricts himself to the plainest terminology of parts and wholes. The whole is apparently the complete unit of analysis: nothing exists outside of it with which he is concerned. The parts exist inside of it, not outside of it, and together they compose it. Kim and Kauffman, though, make things a bit more complicated with their usage of "systems" along with entities and wholes. The whole as a 'system' is used possibly because the author thinks it carries more modern intellectual heft than 'whole'. There might be more to 'system' than there is to 'whole', especially when the 'system' has many intricate and sophisticated parts. Deacon and Ellis prefer 'level', and in two ways: one, because there is a progression from one stage of complexity to another; and two, because there is a demonstrated difference between the properties and powers of the top level and those of the levels beneath the top level. Considering the discrepancy here on the meaning and usage of 'level', one might refer to the former as 'phasal' and the latter as 'ontological'.

Assad and Packard also prefer levels, but they also specifically reference a hierarchy. Hierarchies might have levels, which might be categories of their constituents, but it is certain that the constituents of a hierarchy are *ordered*—at least when confined to a level within a hierarchical structure. There may be an infinite number of levels existing universally beneath the totality of the universe: any arbitrary designation of vertical composition can be called a hierarchy. Nevertheless, one level cannot acquire relative or comparative properties unless a finite hierarchy is established. Because in a given analysis the total number of levels has been determined, a totality arises out of the collection of levels stacked upon each other. The sequential ordering of composition still leads to an ultimate unit of which nothing, no part and no level is lacking. This is our initial definition of whole, and I hope to show that this definition—or, rather, this amalgamation of definitions of wholes—is a marriage of integration (unity) and individuation (distinction) all converging on a singular, *unified* unit of analysis.

1.4.2. Levels and hierarchies

Is there an important distinction between levels, hierarchies, and wholes? When dealing with levels, again, the key adjective is 'comparative'. There must be a comparison between levels: higher to lower, complex to simple, and more-complex to less-complex. Since levels are comparative, two levels cannot stand on equal footing as one must be higher or lower than the other. The analysis must begin somewhere and somehow, so one cannot analyse a higher level and not a) a lower level, and b) an even higher level. In order to analyse, there must be a *highest* level, and the highest level—which might be called *system-level* or level of the whole—is identified in the same way that an object is: it is first individuated as the unit of analysis such that the lower levels are contained compositionally within it and hierarchically below it. By selecting this unit, one has established a base for all measurements and approximations—including those of properties. In order to observe, describe and explain this unit, one necessarily examines its composition.

A 'hierarchy' must eventually reveal itself to be finite as it would otherwise be meaningless, eventually spreading out into a network map of some kind. Because it is finite, a hierarchy is, or is said to be the interpretation of, the observation and recognition of the levels it claims to be composed of. In other words, a hierarchy is not necessarily a discrete, spatiotemporal object in and of itself: it signifies the recognition of a finite pattern and an ordering of importance amongst a finite collection of objects.

If the pattern and order has been recognised, and the hierarchy is limited to a discrete finite collection of determined, stacked levels, then the collection is unified and would therefore be the unit of analysis and the whole.

1.4.2.1. The random grab

Inversely, the lower levels in this ordered collection form part of a receding sequence of hierarchical composition. However, the level terminology can only be used in the context of hierarchies: a hierarchy does not necessitate levels *per se*, though it would normally include them. A hierarchy must be composed of certain things, and since the order is important, levels should be the means by which those things are extrinsically sorted. The hierarchy to begin with results from an intentional selection of a designated sequence of levels, which were themselves initially individuated based on pattern and order recognition. Of course, it is difficult to have any analysis whatsoever without an individuation of an object and the identification and ordering of many objects. It is however unjustified to claim that an object, which may itself be described as having an internal hierarchy, forms part of a large hierarchy of many other objects, without expressing the imposed, ascribed, extrinsic ordering and sorting that has been done—no matter what the ontology. This is to say that 'hierarchy' is not overwhelmingly helpful in explaining emergence, because it is not clear how it can be used outside of the 'random grab' of levels that any hierarchical study attempts to make.

Imagine a person who is reaching into a box filled with small machine parts (from, let us say, a home appliance repair shop) of different shapes, sizes, sources, functions and various other properties. The 'grab' is random because the person cannot see which item she is retrieving until her hand is pulled out of the box. The object acquired is a singular unit of analysis, and any further analysis of it considers only the object, its properties, and its material composition. The object can be examined, and it may become apparent that it has its own subcomponents, which can be divided into levels of complexity or functional importance. The person may differentiate between these levels such that certain components and their functions within the object are considered simple, whilst others are considered more elaborate and interconnected with other components. *But*, the object and its components acquired some of their properties after others: some of them are functional properties of the assemblage, such as how they serve the intention of the object's designer. Other properties, such as of their base chemical components, existed before (and exist now independently of) these functional properties were acquired. Hence, the object—though it is a unit of analysis, and an arguably a whole—only has an extrinsic and conventional hierarchical structure. Whilst it is true that one particularly elaborate subcomponent of the object may itself have the same or similar base chemical components as another less elaborate subcomponent—thus creating the illusion of an intrinsic structure and order of things—there cannot be an essential hierarchy in the order of the object's subcomponents. In plainer terms, one cannot confuse functional hierarchy with intrinsic hierarchy. In an Aristotelian or Thomist sense of the word—though not at all in an Aristotelian application of it—the extrinsic hierarchy is *accidental*, or situational, or observation-dependent (and in most cases subjective).[78]

Metaphysical, intrinsic hierarchies require an observation-independent order and structure: this is to say, an immutable and universal description and explanation of an object's properties (and behaviour: properties mapped over time all being tied to the same object). Some examples of such, historically speaking, are easy to dispute, such as the idea that humanity is the centre of 'creation', and at the top of the pyramid of earthly creatures.[79] Others are not so easy to dispute: a mammal's cognitive and nervous system features seem to inform and even supersede other systems. Is this evidence for an intrinsic hierarchy in mammalian anatomy? This is most likely not the case, but it is difficult because we have trouble defining what 'cognitive and nervous system features' are, for example. We struggle to establish a closed set of parts $\{o_1 \ldots o_n\}$ and properties $\{P_1 \ldots P_n\}$, not to mention the problems resulting from including or excluding the controversial, more volatile parts $\{m_1 \ldots m_n\}_Y$ and *their* properties and sufficient conditions (see §2.8). Inevitably, intrinsic hierarchies must include both intrinsic properties and complete compositional knowledge. This brings us back to §1.0 and §2.8 with the following problems:

[78] One might distinguish between observation-dependent and subjective because the former might be, for example, via unanimous, mass-orchestrated, multi-lateral experimental testing.
[79] This is ultimately a metaphysical claim rooted in, for example, theology and theological bases for claims for free will.

- *(1) Revised*: We know, or have observed, each and every member of the set $\{o_1 \ldots o_n\}$, and they are all physical parts of the whole.

- *(2) Revised*: We have access to all the properties of each and every member of the set $\{o_1 \ldots o_n\}$, and they are all physical properties.

If an object is purported to have an intrinsic hierarchy, and a hierarchy is not an object in and of itself, we should say that the object has intrinsic hierarchical features. Hence, the object is intrinsically hierarchical, and this makes 'hierarchical' an intrinsic property. But, as we have said, intrinsic properties are observation-independent properties: universal, immutable, and lacking any external reference. These properties are not ascribed, but merely discovered, which reveals an alternative ontology, potentially including form or substance. Since this thesis is considering claims in the metaphysics of emergence with an empirically-sound physicalist foundation, I shall not treat claims for intrinsic properties with these alternative ontologies in mind. I will however consider the status of intrinsic properties as requiring, as mentioned in §2.8, a total knowledge of the object's composition and perfect awareness of its properties. Because I consider this to be a flaw or drawback, intrinsic properties are set aside in favour of extrinsic properties, which could potentially include hierarchical and levelled features. Furthermore, these extrinsic properties must, as is the case with all properties, be attached to individuated objects.

Though hierarchies and levels are useful for breaking apart and sorting related objects, which leads to the ascription of patterns and orders, relations between levels can only be explained by comparing properties—which can only be ascribed to objects (or other, lower-level properties, objects, events and conditions). Furthermore, the existence of actual ordered levels in the form of an intrinsic hierarchy can only be based on the declaration of complete causal sequences, which again cannot be confined to the random grab of levels. The utility of hierarchies and levels is this random grab: the object of the 'grab' is itself the unit of analysis and the whole. Is the whole (a) *necessarily* or (b) *instrumentally* integrated and individuated? This is one of the key background questions of the thesis.

1.4.3. Wholes as integration + individuation

A system may have intricate internal processes, functions and sub-aggregates: a machine as a system has parts that interact with each other, and some of these parts themselves may have many delicate inner-workings. There might be a disproportionate amount of potency or "realisability"[80] held by what is really only a fraction of the system: one mechanism within the system might consistently draw the highest proportion of electrical current of all the mechanisms inside. Any examination of the system's internal processes may reveal strange properties, but these micro-level features are not where explanation and analysis begins in the metaphysics of emergence and reduction. The process of investigation begins with the object, the system, and the unit of analysis. It begins with the whole.

The whole is what the metaphysics of emergence requires. It requires an object, to which we ascribe properties, that is complete. It is, from our amalgamated definition of whole, that of which nothing, no part and no level is lacking. This totality is therefore integrated. It is because it is integrated that we can justify, potentially, the object's individuation. In the metaphysics of emergence, reducible and irreducible properties are ascribed to that individuated object. Once individuated, the analysis of strong claims can begin.

The whole is the philosophical confluence of integration and individuation. We could analyse surprising, unique, seemingly unpredictable features of any number of objects, but until the objects are seen as integrated (as the authors above have assumed) and therefore individuated (as the authors above have also assumed), these miscellaneous objects are not wholes as we have initially defined them. Strong claims in the metaphysics of emergence are implicitly or explicitly built upon assumptions of integration and individuation, and are therefore built upon the whole. Chapter II will look at what this confluence is all about.

[80] Kim uses the terms "realizer" and "realizability" in the discussion of emergent causal powers. (1993b, 1999a, 2006a, 2006b)

1.5. Chapter Summary

Weak emergence has not been the source of inspiration for this thesis in the way that strong claims have, themselves involving irreducible, downwardly-causal powers. As has been shown, strong claims all assume a singularly identifiable and unified unit of analysis. Intriguingly, most of the emergence literature assumes the integrity of these objects, such that the object's unity is apparently off the table. If integrity is assumed, where nothing, no part, or no level is lacking, the metaphysics of strong emergence is irrevocably mereological. Emergence, as I see it, needs a whole, and not simply because they are both popularly associated with an object being greater than the sum of its parts.

It should, then, be unsatisfactory to scrutinise strong claims without an engagement with what a whole has historically been or what it potentially could be. More importantly, how can one claim that an object has system-level, self-inflicting, autopoietic causal powers, or that it is merely the cumulative interaction of a finite series of microphysical agents? The discourse requires an integrated and individuated object, and the nature of its integration and individuation must be examined through a mereological lens. By what means is it integrated? What are the relations between parts, or between the parts and the whole? Answering these questions may require a broader worldview, or a subscription to a broader philosophical system and metaphysics—such as those presented in the next Chapter. Nevertheless, scrutinising claims in the metaphysics of emergence has been done, and will continue to be done, in the spirit of this thesis, against a backdrop of what I think is a more empirically-sound physicalism—one that goes beyond the ontological primitives and closed-system linearity of ontological reduction and the confining *règles du jeu* such a position has seemed to impose on the discipline.

This work is willing to bring in a range of subjects to help treat some of the problems in the metaphysics of emergence, and my goal is to make precise, yet deep cuts into this discourse. As a brief recapitulation, I believe this is where we are at this point:

- Properties can only be ascribed to an object, and system or macro-level properties can only be ascribed to all that is the object. Emergent properties are system-level properties, so emergent

> and system-level properties can only be ascribed to all that is the object. That which is all that is the object must be an integrated object, so property ascription to the integrated object is property ascription to all that is the object. Because properties can only be ascribed to objects, and property ascription is to an integrated object, emergent and system-level properties must be ascribed to individuated objects. This is more clearly expressed as saying that these higher-level properties themselves may not merely enable the object's individuation, but that maybe they themselves *aprioristically* individuate the object.

Can an empirically-sound physicalism permit knowledge of either a persisting or instantiated totality? Can it permit knowledge of a complete catalogue of an object's properties? But, perhaps more broadly, must wholes, in a refined physicalism, *necessarily* be integrated and individuated, in order to have system-level and potentially strongly emergent properties? Or, is it ultimately only *instrumentally* integrated and individuated, with our constructions of physical aggregates being merely applicable to our descriptive and explanatory models? These questions will continue to hover over our analysis. More urgently though, Chapter II will introduce the relevant mereological themes to this discussion in the metaphysics of emergence: how is an object to be integrated, or how is a whole made, or, in the plainest of terms: what is a whole?

Chapter II | What is a whole?

Chapter I defended the claim that strong forms of emergence in particular must involve a discussion of parts and wholes, where a distinction between wholes, systems, entities, objects and units is unnecessary. Because novelty, irreducibility and downward causation together require comparisons between properties in and of the same unit of analysis, a conceptualisation of *whole* is central to a conceptualisation of (at least strong forms of) emergence. More specifically, claims in the metaphysics of emergence have thus far been based on assumptions of a given object's integration and individuation, and in doing so, as discussed in Chapter I, make the metaphysics of emergence extend into both ontology and epistemology. By connecting the metaphysics of emergence to mereology and the various traditions of part-whole relations, I believe we can address its claims more directly.

2.1. Introduction to mereology

Mereology, the study of part-whole relations, is at the core of the classical problem of the one and the many. How can something that has many parts at the same time be one? Within western philosophy, there is virtually no period in history since the Pre-Socratics that fails to comment on this general question. Among others, Heraclitus, Democritus, Plato and Aristotle all make substantial contributions along with the Scholastic authors Raymond Lull, Peter of Spain, Boethius and Aquinas. Leibniz borrows some of the 'combinatorial' mereology from Lull and some other terminology from Aquinas, including a helpful discussion of the scholastic invention of the *integral whole*—a device which will prove useful to this thesis. Brentano and Husserl follow Aristotelian mereology in many regards, often building on it to incorporate it into the wider phenomenological programme. All of these authors specifically pay detailed attention to a theory of parts and wholes *in those terms*, thus making the theory *mereological* in a stricter sense.[81]

The selections are broken into early and late thought, treated in both sections chronologically. I start with classical atomism, Plato, and Aristotle, and then finish §*1.0* with some of the helpful medieval contributions. §*2.0* initially looks at Leibniz and Husserl, and then moves on to the mereological calculus of individuals, set theory, and other twentieth century developments. Throughout this brief survey, we will keep in mind four themes to unify the reading, and not necessarily in the same sequence or in logical progression:

- (1) *the status of the whole's unity*

In Chapter I, I mentioned that one of my key concerns in the metaphysics of emergence was what I perceived to be the assumption that questions of integration could either be ignored or could be placed after questions of individuation. I have tried to emphasise that we should be confident that objects are integrated before discussing the role of individuals. This first theme will help organise our mereological options so that we know we can actually ascribe system-level or macro-level (and potentially emergent) properties to the individuated and compositionally-integrated object and potential whole.

- (2) *the ontological categorisation of both parts and whole*

Some mereologies will include more than matter and force particles, and when they do, an additional category is required. In some cases, for example, the whole consists of the parts all being physical and some aspect at the level of the whole being non-physical. If some component of the whole at any level does not match the categorisation of the rest, it will need to be identified—especially if claims for the non-physical are central to the metaphysical system in which it is found.

[81] Boethius in *Theological Tractates* states that every axiomatised metaphysical doctrine has a mereology as its component. Leibniz in *Ars Combinatoria* defines metaphysics as the doctrine of the totality and the parts if considered per se. Kant in *Monadologia Physica* mentions the need for a mereological theory in order to deal with the theme of the constitution of *sein*: "bodies... are composed of parts and it is certainly not an easy task to establish clearly in what way they are composed of these." (1910, 476)

- (3) *dependency relation between parts and whole*

This is potentially a combination of all four of our themes, but initially it simply calls for an analysis of the various ways the higher level influences the lower level and vice-versa. It may come down to asking whether the whole needs its specific or proper parts to exist, or whether the parts need the whole to instantiate them. I will come back to this tenuous dichotomy later in the Chapter.

- (4) *relevance of order and structure of parts to whole*

Can an object or whole survive the addition, subtraction, or re-arrangement of its parts? Must there be an intrinsic order and necessary composition and configuration in order for object X to be whole W? This dimension can have at least two different paths. The first concerns persistence (addressed in Chapter IV), and the other concerns the flexibility of our definition of composition. Does it matter whether the lower levels are dynamic and diachronic, or fluid and porous? In some systems, the order and structure matter, and the existence of the whole depends on a fixed arrangement of the parts.

As we progress through these contributions, I will use these themes with the mereologies and try to identify what our options are in determining the nature of the whole. From there, I hope to establish criteria for strongly emergent or reductionist property ascription that can then be scrutinised. The same obstacles remain: are we fundamentally restricted by (even a refined) physicalism? Can we overcome the conflation of the epistemological and the ontological? And, can we justify our integration and eventual individuation, should such a physicalism allow it? We start with Pre-Socratic materialism.

2.1.1. Atomism

As a stark contrast with the rest of classical part-whole discussions, the matter-and-void atomism of Leucippus and Democritus, and later by Epicurus and Lucretius, presents a thoroughly materialist ontology. In this way, a look at these doctrines should prove to have some connection with contemporary physicalism as well as with whole determination.

Atomism can be *prima facie* semantically interchanged with reduction, and more specifically, ontological (metaphysical) reduction. 'Uncuttable' units of analysis are discoverable for any form of atomism, and this is the start and end of the school. There have been numerous sects and approaches based on of the idea (e.g. logical atomism), but all of them have been met with strong opposition simply because indivisibility is a 'difficult' claim to substantiate.[82] Perfectly isolating the tiniest particle in contemporary particle physics is impossible, as we, for example, lack the ability to distinguish particles from waves under a vast range of conditions: we do not know what the unit of analysis might be, let alone look inside it. Our theoretical mathematics do not make sense when going smaller than 10^{-35}m: as such, perhaps we should not claim that we have mastered the physics of 10^{-34}m but not of 10^{-36}m.

The core principles of classical atomism are easy to recapitulate, but difficult to comprehend. 'The Void', or "what there is not", is infinite empty space, in which nothing exists. The only other things in existence are "what there is", or 'the Full': an infinite number of atoms—inconceivably, imperceptibly small, rock-like objects that move in the infinite empty space whilst occasionally colliding with each other. (Bailey, 155) The Full needs The Void, and vice-versa, and both are equally real. As Aristotle tries to rebuke in *Physics* (VI) and directly address in *Metaphysics* (Alpha, 985b), the Atomists have the unenviable task of dealing with infinite regression. How can something exist if it is infinitely divisible, whilst at the same time be made of indivisible units? It seemingly must be infinitely divisible, because the totality of the atoms has no limit, and the infinite atoms are indefinitely swirling, crashing and falling—creating and destroying aggregates at imperceptible speeds and durations.

[82] Not being discussed by any of these groups is something which proves to be an even greater empirical threat: ontological dependence. Can we demonstrate that whatever smallest possible unit of analysis in question exists in and of itself? Aside from string theory, what reasoning for this is available? Furthermore, indivisibility in atomism is in its applications accompanied by claims of isolation, static entities, objectivity, logical atoms, observation-independence, and universal and immutable truths. Can any of these things ever be substantiated?

Despite most of the basic physics of these early atomist thinkers being obviously superseded by modern (e.g. sub-nuclear) models—as most of classical physical investigation has been—the standard physicalist may not find much wrong at all with: 'all agency is nothing but the cold, brute, accidental force of chunks of matter colliding with each other.' In fact, a modern physicalism could be based on such a statement of accidental, aggregative collision, as generically opposed to the supposition of unities, essences, souls, or mind stuff. The problem that we should have with the Atomists, I believe, just like with standard physicalism, was whether or not we could confidently say that there *were* primitive chunks to base the aggregates off of. The assumption of a universal bottom-level is certainly one that is not accompanied by observational evidence, in any ontology. And, in the days of Leucippus, it would of course have been mere speculation.[83] Whilst it is enormously helpful to consider micro-level agency, particularly in light of what we know about the microscopic world, classical atomism is among the chief contributors to hasty, bottomed-out whole ascription within standard physicalism.

Are wholes possible, let alone relevant, for classical atomists? If all things are compounded atoms moving in the void, why even bother talking about wholes? If all physical structure is not infinitely complex, is it not possible to conceive of things from which nothing is missing? Furthermore, the smallest possible unit of analysis can have no parts—i.e. a mereological simple. If it has no parts, it probably cannot be a whole. But, it is on this question where differences seem to arise between the Leucippus-Democritus doctrines—much of which is learned via Aristotle—and those of Epicurus-Lucretius.[84] We have no reason to believe that the former set have anything specific to say about 'whole' or 'unity', whilst Epicurus and his Roman expositor two-hundred years after him make occasional references to souls, human bodies, and minds.[85] The former thought that growth was aggregation, where particles non-teleologically came

[83] Though, this is what metaphysics may very well need to be. If there were an accompanying observation, we would be scrambling to make sense of that observation, occasionally with the help of speculative future observations. Throughout the process, we would therefore be bridging observational gaps with language and meaning, a basic prerequisite of metaphysics.
[84] One of the notable contributions here is that of the atomic swerve, or *clinamen*, which allows for indeterminism in the universe—including, for Lucretius, free will.
[85] Some of these comments are made in a letter to Herodotus discussing the fear of death, whilst others are included in *De Rerum Natura*. In this same letter, there are some statements that hint at the whole as being nominally and parsimoniously real. Discussing how these later Atomists might be construed as nominalists is outside the scope of this chapter and thesis.

together. Decay and disaggregation—where particles came apart from one another after having been somehow bonded—eventually led to, in the case of living things, death. The latter appear to countenance some notion of unification of mind and body into soul, but these references are scattered. If there is unification, the soul is a whole unlike the aggregates of mountains and trees.

Let us be flexible with the Epicurean branch of atomism and conclude this section with what is perhaps a selective interpretation on this issue. Atomism does not recognise the need for wholes qua unities, because matter collides with itself in the void arbitrarily. A crude summation might be that—when particles crash together in their fall through the void—they can sustain this aggregation long enough to give the impression of structure and/or unification. Using aggregation as the weakest possible *description* of whole, atomists might find use in a *concept* of whole. Now, we will see our four themes. (1) The atomistic whole is only temporarily or extrinsically unified in the form of aggregates. Epicurus might be interpreted as holding that the soul was a combination of mind and body, and possibly a unified whole. (2) Since atoms are imperceptible, all objects of perception are aggregates, and therefore of the same ontological category as their pieces. (3) The whole depends entirely on its parts for its existence and individuation—including its duration. (4) The order, configuration and structure of the parts, being ontologically primary, matter, as they alone determine the nature of the aggregate or whole.

2.1.2. Plato

In Plato's later dialogues, he tells us that collection and division provide us with a way to understand the relationships between some unity and some plurality (*Phaedrus* 265d-266b, and *Philebus* 16c-17a). Whilst this may have epistemological implications, it *must* have ontological implications. As we saw in Chapter I, reduction in this general sense is step one in all investigation. What kind of ontological priority do we grant to these kinds of objects, and are they integrated, individuated wholes?

Verity Harte (2002) and Eric Brown (2004) have both written that Plato can be interpreted as seeing wholes as both ontologically *innocent* (having no special ontological status) as well as ontologically *creative*,[86] and they find places where both positions can be defended. In some passages (notably in *Theaetetus*), Plato seems to suggest that a whole *just is* a "many", and hence, composition *is* identity (Harte, 2002, 43). In this sense, an object's properties are exclusively additive, and the whole is nothing more than the sum of its parts. This rejects any notion of unity in the "many" even though we can loosely collect the many without any acknowledgment of singularity. However, in *Parmenides* and especially in the *Sophist*, *Philebus*, and *Timaeus*, Plato argues that a whole is a genuine, complex individual, and an entity created by the composition of its parts. Harte attributes to Plato the view that structure (or configuration) is the crucial element in whole ascription—an attractive and evidenced position—and this leads us to the notion that wholes are "contentful structures." (178) I have found it a bit challenging as to how these two obviously contrasting views might be reconciled, especially when ascribing "content" is easily confused. Because of these apparent inconsistencies, Plato will be much less easily categorised than Aristotle and the Scholastics.

Plato's conception of a whole expressed by Socrates in *Theaetetus* (205a4–7) as "that from which nothing is absent" is intriguingly different from the characterization of a whole, at *Parmenides* 137c7–8, as "that from which no *part* is lacking" (which is subsequently repeated by Aristotle in *Metaphysics* to describe the "all"—which is also used by the Scholastics as a *totum*). Even without an Aristotelian essence, wholes that have nothing absent must also be unities, otherwise the distinction that nothing is missing would be irrelevant. The question "from what is nothing missing?" is answered with "something which must be unified." So, in *Theaetetus* we see a case for unities, though we are not sure how they are unified, as many unities—such as syllables in 205a—do not have forms or essences. *Parmenides*, which is a source for substantial mereological analysis, emphasises the role of proper parts:

[86] Let us call this Platonic mereology 'Version B'. Version A is what follows as 'contentful structures'.

[137c]
SOC: "Well then, if the one exists, the one cannot be many, can it?"
"No, of course not."
"Then there can be no parts of it, nor can it be a whole."
"How is that?"
"The part surely is part of a whole."
"Yes."
"And what is the whole? Is not a whole that of which no part is wanting?"
[137d] "Certainly."
"Then in both cases the one would consist of parts, being a whole and having parts."
"Inevitably."
"Then in both cases the one would be many, not one."
"True."
"Yet it must be not many, but one."
"Yes."
"Then the one, if it is to be one, will not be a whole and will not have parts."
"No."
"And if it has no parts, it can have no beginning, or middle, or end, for those would be parts of it?"
"Quite right."
"Beginning and end are, however, the limits of everything."
"Of course."
"Then the one, if it has neither beginning nor end, is unlimited."
"Yes, it is unlimited."
"And it is without form, [137e] for it partakes neither of the round nor of the straight."
"How so?"
"The round, of course, is that of which the extremes are everywhere equally distant from the centre."
"Yes."
"And the straight, again, is that of which the middle is in the nearest line between the two extremes."
"It is."
"Then the one would have parts and would be many, whether it partook of straight or of round form."
"Certainly."
"Then it is neither straight nor round, since it has no parts."
[138a] "Right."

(*Parmenides* 137-138)

The argument appears to be that because something without parts is infinite, all finite things must have parts. Unities, wholes, and *the all* must have parts, but unities and wholes are not missing any parts. Since in this chapter we are looking for a range of answers on the definition of a whole, we can say that from *Parmenides* a whole seems to be something which has parts, of which none are missing. Unfortunately, this can collapse quite easily, and in two ways. One, how can something be lacking a part, if we do not know what parts are missing? Two, if we know which parts are missing, then surely we are ascribing such incompleteness based on an ascribed completeness. How can we ever be certain of completeness? The syllable is incomplete without a letter, as in *Thaetetus* 205, but that presupposes a global existence of *syllable*. Since language determines what "syllable" is, language then also determines completeness of *syllable*.

With Platonic forms looming in the discussion, are wholes the "all" with the form of one type of "all"? In other words, is a red apple a whole because it shares some of the form of "red" and some of the form of "apple"? This might be initially appealing most especially in the context of real numbers: are two red apples, each sharing some of the form of "red" and "apple" considered two because they both share in the same forms, and are thus countable and of a set? The number two is an abstract or intensional whole arguably because the whole is composed of two things sharing the same form. But there are at least two problems with this. First, we are not to assume that the two objects sharing one form are sharing all forms, and thus are not necessarily identical. This means that when we count, we are assuming that all members of a set are individuated and identical, even if they are numbers themselves—for the numbers must correspond to extensional (spatially-extended, empirical, spatiotemporal, 3D or 4D, etc.) or intensional (mental or representational) objects. Second, in order for "two" to be a whole, it must be more than a set or aggregate or *the all*. But for something to share in the form "two", it must be paired with an identical object, thus forming a set. Let us assume that Platonic forms are not in the discussion of what might be a whole.

Appraising Platonic wholes becomes a bit less tenuous later in *Thaetetus*, where he concludes that a whole cannot be composed exclusively of parts, because then it would be the same as all of its parts. Plato appears to make up his mind, not wanting to designate *the all* as a whole. This leads us back to Harte's "composition is identity" and implies that we can only know what a whole is—or which things are wholes—if we know whether and how the whole in question survives the addition and subtraction of certain parts. The "survivability" of the whole in the face of the addition and subtraction of its parts—where it allows us to determine whether something is a whole or not—is of enormous importance to mereology (and by extension, emergence) for several reasons. (A) This may be an acknowledgment of either temporal parts, temporality, or at least diachronic composition, something that many philosophers have still not accepted or have chosen not to treat. (B) This may be an acknowledgment of what we will come to call *integral wholes*: wholes which are composed of things that they *necessarily need* to survive (dependent parts, e.g. a human heart to a human) as well as things that they *do not necessarily need* to survive (independent parts, e.g.

molecule X to a human). (C) It leads us to believe that we will not know which objects of perception are wholes without observation—a stunning revelation of a quasi-empirical and borderline-nominalist implication in Platonic mereology. If we are boldly selective, and we consider Platonic wholes as being "contentful structures"—some of which, of course, are unities, with, for example, souls—then we can consider a whole range of real-world spatiotemporal objects under contemporary naturally-scientific scrutiny.

Before finishing with Platonic mereology, I believe it is important to mention briefly why there is a potentially-important difference between a whole as (a) 'that from which *nothing* is lacking' and (b) 'that from which *no part* is lacking', because this question, at least in some sense, will play a role in later sections and chapter. When we say that a whole *has parts*, we could also mean that it has *nothing but those parts*, whether or not a whole is doing something more than what we would expect of its parts. It is possible that a *candidate* emergent object qua whole has irreducible and unexplainable macro-level powers or properties, where we exclusively know and treat its parts to verify it. In other words, when we choose (b) as a definition of whole, the only thing used to assess the claim of supersummativity is its parts. The definition in (a) evidently does not require exclusively parts, which means that (b) does not need to be empirical. Much more will be presented on this issue as we move forward.

Summarily, Platonic mereology can be construed in different ways, but let us look at our four themes. (1) Plato considers wholes as systems, but not always as unities. With those that are unities, some have metaphysical unifying essences (loosely used), such as souls. Many things in the world are not unified, but share in some forms, which are tenuously wholes (though not really relevant to our physicalist base). (2) Unities have at least two different ontological categories, whilst aggregates have one (and he is unclear as to whether these should really be called wholes). This is to say, some wholes share the same ontological category as their parts, whilst others do not. (3) In unities, the existence of the parts does not always depend on the existence of the whole, but many "contentful structures" have some parts whose loss the whole can survive. (4) The configuration of the parts matters in some unities, but not all, since we can acknowledge that some parts can be lost. The lack of coherence in his mereology might arguably place his entire philosophical system in question, though we might postulate that Plato's mereological contributions resonate more with a refined physicalism than those of our next author.

2.1.3. Aristotle [87]

Aristotle, the author of the most celebrated classical mereology overall, even if simply as a result of the quantity of his writing on relevant issues, states the problem of the one and the many in *Physics* with the intention of addressing it later in *Metaphysics*:[88]

> There is, indeed, a difficulty about part and whole, perhaps not relevant to the present argument, yet deserving consideration on its own account—namely, whether the part and the whole are one or more than one, and in what way they can be one or many, and, if they are more than one, in what way they are more than one. (*Physics* I.2, 185b11-14)

Through the central books of the *Metaphysics* the author is concerned with the distinction between a compound[89] thing which is either a whole or a heap (*soros*). The whole is that thing which contains more than its elements (or matter) which seems to account for its unity that we observe in that thing, whilst a heap remains a compound object lacking such unity. As Koslicki writes, Aristotle's position is that the means for unity of the parts, a "something else", is its form or essence[90], and thus a whole (or unified compound object) has this form and a heap does not.

Aristotelian heaps are the simple aggregate with all parts being separable. He cites the example of a pile of sand, where the parts can exist without the 'whole'[91] (itself contingent or accidental). The heap or "the all" is a set of particles or pieces (and not

---

[87] Aristotle will be treated here in extensional terms. This is largely due to the fact that an intensional mereological treatment—conceptual, and not material parts—is scarce. If it were to be pursued, it might be, as Libardi hints, in the area of the Aristotelian notion of definition.
[88] He sets up the problem in *Physics*, expounds upon it in *Metaphysics*, and raises broader "Problem of the One and the Many" questions in many other texts (including *Poetics*, *Topics* and *De Partibus Animalium*).
[89] The word "compound" here presents a major opportunity for digression, but cannot be discussed in detail. The key understanding is self-evident: having many discernible parts.
[90] Koslicki uses the terms interchangeably when referencing the "something else". I see no reason to deviate from this choice, considering the brevity of this present treatment and the rather irrelevant distinction that could be conjectured by it.
[91] I add inverted commas here to indicate that the reader is reminded that this is precisely the antithesis of 'whole' as shown above and below the sentence.

necessarily 'parts') with no particular order or configuration, lacking any intrinsic structure. Conversely, on the subject of the "the one" Massimo Libardi writes:

> It [Aristotle's description of the nature of the whole] implies an order of the parts whose modification affects the whole itself. The whole therefore presents an internal structure and it is characterised by the fact that if certain parts are removed, then it ceases to exist. In other words, it is defined by the fact that some parts are separable and others are not. An example is the human body where some parts (such as the head) cannot be separated, while others (such as the fingers) can be amputated without the person dying. (1994, 16)

As we are beginning to see here with Aristotle, mereology is not necessarily exclusively about object composition and the relations between parts and between parts to whole. The implications of these things make the stakes of mereology very high because they must also relate to essence and change. To ask what a whole is, is also to ask how a whole comes to be: they both depend on a foundation for existence. Do wholes simply appear, spontaneously, or, as it were, *emerge*? Consider a common object, like a car, and how the question of what the car is (and what parts it has) is actually dependent on the question of how the whole comes to be. The question of *what the car is* at this moment, is different from, and yet dependent on, the question of *how it exists* as such.

Figure 2.1.3.1: Shuffling V. Instantiating

<table>
<tr><th>Statement</th><th>t_1</th><th>t_2</th></tr>
<tr><td>Shuffling Parts
(A) To change the parts is to change the whole</td><td rowspan="2">Integrated, individuated car (necessarily) has wheels (and tyres)</td><td>• No car
• Car minus wheels (and tyres) is unorganised heap
• Wheels (and tyres) removed and set aside</td></tr>
<tr><td>Instantiating Parts
(B) The whole instantiates the parts</td><td>• No car
• Car minus wheels (and tyres) is unorganised heap
• Wheels (and tyres) are now just 'steel and rubber', or just 'matter and force particles'</td></tr>
</table>

Whilst on the surface they appear only to differ in the level of initial analysis—as in, from the perspective of one level or the other—they might involve a fundamental difference in ontology. Modifying the parts like (A) in some way prompts a vision of removing the wheels (and tyres) from a car. The object ceases to be a car (at least from a functionalist perspective), but the parts are to be considered as *discrete*—like removing a brick from a wall or a shoelace from a shoe. It seems to me that this is not the same as (B), which really says that whatever the parts are at t_1, they do not exist at t_2. This could mean that the steel-rubber aggregate that had been a wheel (and tyre) at t_1 can no longer be considered a wheel (and tyre) at t_2. It could also mean that the unified, structured object of matter and force particles we had called a wheel (and tyre) at t_1 is at t_2 simply an unorganised, unstructured heap of elements. In other words, in both cases, the car ceases to exist at t_2. In (B) however, the wheels (and tyres) also cease to exist. This suggests that the whole *instantiates* the matter-and-force aggregate qua parts. The Philosopher does not, however, leave this problem completely untreated.

The ideas of separability, actuality and potentiality are crucial to Aristotle's mereology. Material parts are not separable in so far as once they have been separated they are no longer parts but potencies. Libardi continues: "the hand is no longer a hand if separated from the body and in the same way neither is the eye if separated from the face." (1994, 37) This means that the existence of parts qua parts is dependent on the existence of the whole. This is not to say that the part which has been separated from the whole (the "part" still being a *thing* and extensional, or spatially-extended, and still composed of matter and force) no longer exists. Rather, in this mereology, it ceases to be a part. Thus, it is not an ontological question *per se*, but purely a mereological one. One example is that "the amputated finger is still something." (37) But, there are two challenges that Aristotle does not adequately meet. First, what happens to the part after *kolobon* (mutilation)? Does it have any structure at all, or does it become a heap in all cases? Second, if the whole survives *kolobon*, how should we categorise those parts that have evidently little influence on the whole (or at least its essence) before they were removed? They are parts, and yet are not necessary. These are the issues that the Scholastics[92] will try to address.

Very simply, the whole determines the parts, and thus the scope is top-down. The following passages from the Zeta book reinforce this notion:

> Now since that which is composed of something in such a way that the whole is a unity; not as an aggregate is a unity, but as a syllable is—the syllable is not the letters, nor is BA the same as B and A; nor is flesh fire and earth; because after dissolution the compounds, e.g. flesh or the syllable, no longer exist; but the letters exist, and so do fire and earth. Therefore the syllable is some particular thing; not merely the letters, vowel and consonant, but something else besides. And flesh is not merely fire and earth, or hot and cold, but something else besides. Since then this something else must be either an element or composed of elements, [20] (a) if it is an element, the same argument applies again; for flesh will be composed of this and fire and earth, and again of another element, so that there will be an infinite regression. And (b) if it is composed of elements, clearly it is composed not of one (otherwise it will itself be that element) but of several; so that we shall use the same argument in this case as about the flesh or the syllable.
> (*Metaphysics* Z.17, 1041b11-24*)* [93]

Now, if the letter is still a letter without the syllable, but the hand is not a hand without the body, we have two different takes on the shuffling v. instantiation issue. Nevertheless, here above we see a need to account for the 'something else', and how for him ontological reduction (i.e. atomism) seems to prove inadequate because it leads to infinite regression. Infinite regression would seem to deny *de facto* integration and individuation, and, therefore, amongst other outcomes, deny explanations in causal terms (let alone identification of causal relations). Continuing on this point, he writes:

> It would seem, however, that this "something else" is something that is not an element, but is the cause that this matter is flesh and that matter a syllable, and similarly in other cases. And this is the substance of each thing, for it is the primary cause of its existence. And since, although some things are not substances, all substances are

[92] The notion of the whole for Aquinas (to be discussed later in greater detail), who builds on Aristotle, requires that every part of the whole exists if and only if the whole exists. This also means that what applies wholly to the whole also applies to each of its parts. It should be noted that Aquinas also adds another rule: that no separate part shares the form of the whole.
[93] In the middle of this passage is essentially the admission that the form (a "hylomorphic" or matter-form compound) may be the substance and the cause of its existence, which is of course altogether confusing. This confusion is the subject of Koslicki's article.

> constituted in accordance with and by nature, substance would seem to be this "nature," which is not an element but a principle. An element is that which is present as matter in a thing, and into which the thing is divided; e.g., A and B are the elements of the syllable. (*Metaphysics* Z.17, 1041b25-33*)*

This is the key passage: something in the whole that is not physical makes a heap a whole, and it is what integrates and individuates. The statement that this substance is “not an element, but a principle” seems to suggest that no observation could detect that which integrates and individuates. Is this where physicalist analysis stops, or where the dialogue between Aristotelian or essentialist systems and physicalism ends? On the surface it may seem so, but more will come on this later in the work. In short though, these passages above show us how Aristotle *integrates*.

In *Metaphysics*, he is concerned with what makes the unified whole more than just the sum of its parts. In *Physics*, he emphasises the closed nature of the unity:

> What is true of each particular is true of the whole as such-the whole is that of which nothing is outside. On the other hand that from which something is absent and outside, however small that may be, is not 'all'. 'Whole' and 'complete' are either quite identical or closely akin. Nothing is complete [teleion] which has no end [telos]; and the end is a limit. *(Physics,* Book III, Chapter 6, 206*)*

The whole for Aristotle must have limits and boundaries, and those boundaries are ultimately determined by what we should properly call its essence (because ‘essence’ encapsulates the functions of ‘form’ and ‘substance’ regardless of object). This is how Aristotle *individuates*. He individuates because he has shown why the whole is *integrated*. One might even say that claims for Aristotelian *individuation* emerge from claims for Aristotelian *integration*.

This is how we can summarise his mereology: (1) A unified whole (‘the one’) has unity through its form/substance (essence), whilst the non-unified heap has no substance (nothing “rests under” the elements) or a ‘something else’ (only ‘accident’). (2) Because of a unity’s substance and therefore essence, the parts and the whole belong to different ontological categories. (3) The whole’s essence determines what parts the whole can have. As such, parts of *the one* cannot be separated from it. (4) Following from (2) and (3), the order and structural configuration of the parts are both relevant.

2.1.4. Notable Scholastic contributions

Most medieval mereology uses Aristotelian concepts like substance and essence, and discrete and continuous quantities (*Categories*) on top of the standard *collection* and *division*. Most if not all European philosophy during this period was obviously Christian, so logicians like Peter of Spain[94] and Thomas Aquinas were first and foremost theologians.[95] As a result, reconciling Aristotelian mereology with, for example, Abrahamic treatments of souls, turns out to be not only historically necessary[96] but also an easier task than one might think in the context of wholes: there must be something more than the parts, because of its structure and form, and also because it is confirmed by *revelation*.[97]

2.1.4.1. The origin of integral wholes

Aristotle presents two possibilities of *whole*: the heap (the all) and the one. The latter is construed by medieval authors as a *totum essentiale*: an entity which has parts, features and properties specific or essential to a *topic* (for Boethius and Peter of Spain), or class of things. As such, for man to be a rational animal, which revelation instructs us that he is, he must have free will, imagination, and, for some, a sense of humour (what we might today call 'personality'). These are essential parts of an essential whole, and the rational parts of the rational animal. The normal requirements for the animal side of the rational animal would be having two hands, two feet, and a host of other anatomical and, by extension, functional-teleological (built or made to

[94] The terminology of *totum essentiale* and *totum integrale* dates back to Peter of Spain, *Summa Logicales* 5, 14-23. Additional mereological discussions are found in his *Tractatus*.
[95] Boethius (*De Divisione*, *De differentiis topicis*, *In Ciceronis Topica*) is perhaps not primarily a theologian, but certainly works off of Aristotle and others, and mixes plenty of theology with his metaphysics and logic.
[96] In other words, philosophy in this period is expected to at least be theologically-conscious for essentially political reasons.
[97] These tenets all combine to help explain 1) that thing which makes the whole more than a heap (e.g. the soul), and complement 2) the *prima causa*.

serve an end) parts to complement the rationality. However, one can have but one hand and still be considered rational; so, the properties essential to one class do not all serve the same end. It is through this confusion that the medieval thinkers considered man—an essential whole with some properties not essential to *human-ness*—as another kind of whole: the *totum integrale*.[98] [99]

What does it mean if a man has parts that are not essential to him being a man? First, it implies that man cannot be a perfect essential whole, which is something unable to survive Aristotle's mutilation (*kolobon*). Man can survive the loss of some of his parts—though by no means all of them—because there must be some parts that, when removed, prevent the man from being a man. One obvious example is a soul, though even medieval philosophers would have struggled with the idea that one part of creation that had once been granted a soul has since lost it or has had it stolen. The more intriguing case is one where a man is mortally wounded in the cranial region, suffering irrevocable damage to his cognitive faculties.[100] By this principle, he is no longer a man as he no longer has decision-making abilities or the ability to act on trivial desires—though he presumably still has his soul. The soul, for these thinkers, is what Aristotle would have called, for humans, substance, and—in the context of mereology—what we should encapsulate into essence—as it binds the parts together into a unity from which nothing is missing.[101] [102] This, then, is the crucial point. Some wholes, which are evidently more complex than a heap of sand, can survive the loss of some parts (which are called inessential), but not others (which are called essential).

---

[98] Notably the *totum integrale* is discussed in *On Truth*, Question 25, Article 5 (1954).

[99] Aquinas also presents (*Summa* I, 78, 1) another kind of whole, the *totum potestavitum*, where all parts and potencies are exercised at full potential. One can imagine a celebate monk striving in his life of prayer and adoration to achieve some sort of spiritual perfection—though such *potestavitum* can surely only ever be used to describe god in western theology. This is probably a nod to Augustinian writings, as it is unlikely to be a nod to Platonic ones—presumably the pre-Christian origins of such an idea.

[100] Aquinas would certainly say that his soul is intact, but his will and imagination have been damaged or altered, and as such, judgment as to whether or not he is a man is up to god.

[101] We see now the importance of the confusion found in Plato between the nature of a whole being "that from which nothing is missing" and "that from which no part is missing". The difference is astonishingly important, as, if we concede the existence of occasional non-essential parts, the need for integral wholes is emphasised.

[102] Even Aquinas concedes (*Summa* I, 76, 6) that the soul can be considered the *form* of man's body. This is the *unifying* element required in anything claiming to be more than the sum of the (physical) parts.

An integral whole is that which has inessential as well as essential parts, where essence is determined by external propositions. This is to say, for most of the Scholastics, the essence of 'man' is very clearly stated in Catholic doctrine, as compiled largely by Aquinas.[103] Other propositions might include the predetermination of the essence of an object through its membership in a class: the essence of a particular frog F is known and determined by its membership in the category 'frog'. F's essence is based around 'what frogs do' and 'what frogs are'. The implications are, of course, tenuous. The class 'human' in contemporary or at least non-Christian terms is seemingly impossible to set aside, so establishing an essence of man is by extension gravely tenuous. But, we are still working with that which binds the parts, so how can we make sense of 'whole' that has no essence? We often still turn to the soul, either in universal terms—as theologians do via creation—and particular terms—which other forms of contemporary spirituality might use without any necessary statement on how that or any particular soul came to be. The soul for the former is something all humans have in common given by god through creation. The soul for the latter is something that simply exists, by a range of hypothesised means, and is irreducible, non-extensional (non-spatially-extended), and possibly even non-physical and non-temporal.[104] In both cases, the soul acts as the *form* to bind the parts, thereby creating some kind of unity with structure or configuration.

If one believes in the existence of the soul, man is an integral whole because he possesses:

- ❖ *(1) a few parts and properties—most of which are substantial—that are necessary for membership in 'man', such as:*
    - *a) soul*
    - *b) volition, imagination*
    - *c) heart, brain*
- ❖ *(2) many parts and properties—some of which are accidental—that are not necessary for membership in 'man', such as those not explicitly essential as, or those accidental to:*
    - *a) 'man-ness' (e.g. brown hair)*

---

[103] The essence of man, for standard church teaching, which is relevant to the present section, would roughly be described as follows: Man is the part of creation with an individual, eternal soul and free will, capable of glorifying and serving the creator in this life and the next, and eligible for the bestowing of grace to improve the conditions for achieving this end ideally, though not necessarily through the institutions and sacraments of the one, true Church.
[104] Another feature to include might be irreproducibility, which I would heartily endorse—though for different reasons and in much more than the present context.

b) 'living' (e.g. fully-functioning hand)
c) 'enduring' (e.g. chemicals, molecules or other objects originating outside the body that are not present at all times of 'living')

The soul of course is not a focal point of our present investigation, and integral wholes would not obviously require them *a priori*. An integral whole simply has parts that are both dependent and independent of the whole. What does this mean? As previously mentioned, the integral whole will play a vital role in this and other chapters of the present work, and so I will give a thorough treatment eventually in contemporaneous terms. For the medieval thinkers though, we can see the shift from the black-and-white dichotomy of heaps versus wholes to the acknowledgment of essential wholes that have inessential parts.

2.1.4.2. Quiddity

There is one other concept from this period that should be mentioned, and, like the *totum integrale*, the notion of *quiddity* ('whatness') is one that is used by multiple Scholastic authors. It is, at its core, an attempt to refine, or at least build upon, Aristotelian essence. I will draw exclusively from Aquinas to treat this term, who himself has the fortune of drawing on his Scholastic predecessors like Boethius (and even Avicenna). Because essential wholes needed to be determined by the essence of the thing, and essence superseded Aristotelian form and substance, quiddity is introduced to have a striking similarity with essence—and yet with a very important distinction:

> Because the definition telling what a thing is signifies that by which a thing is located in its genus or species, philosophers have substituted the term 'quiddity' for the term 'essence'. The Philosopher frequently calls this 'what something was to be'; that is to say, that which makes a thing to be what it is. It is also called 'form', because form signifies the determination of each thing, as Avicenna says. Another term used for this is 'nature', using 'nature' in the first of the four senses enumerated by Boethius. In this sense anything is called a nature which the intellect can grasp in any way; for a thing is intelligible only through its definition and essence. That is why the Philosopher, too, says that every substance is a nature. The term 'nature' in this sense seems to mean the essence of a thing as directed to its specific operation, for no reality lacks its specific operation. The term 'quiddity' is derived from what is signified by the definition, while 'essence' is used because through it, and in it, that which is has being.
> (*On Being and Essence*, Chapter I §4, emphasis added)

From this passage it seems to me should come a deep appreciation for the term 'quiddity', because it may bridge the gap between knowledge of the essence of a thing and knowledge of the definition of a thing. If in any further mereological analysis in the metaphysics of emergence we should feel more comfortable with discussing 'whatness' instead of essence or substance (which may require a wider subscription to given metaphysical system) to justify our integration and individuation (which is seemingly necessary to investigate), we should keep quiddity in mind at least as a potential placeholder for *that which enables wholeness.* If this is the case, integral wholes with discernible quiddities are profoundly relevant for the metaphysics of emergence—and both concepts derive from this period.

Our themes for the Scholastics will to some degree mirror Aristotle. (1) Wholes that are aggregates have accidental form, whilst wholes that are unities (*totum essentiale* or *totum integrale*) possess a substantial or essential form—the soul, in the case of humans. (2) The *totum essentiale* and the *totum integrale* have a substance or form, and therefore essence, so they have two different ontological categories. (3) *Totum essentiale* cannot survive the loss of one of its parts because its essence determines its composition—the whole determines its parts. *Totum integrale* can survive the loss of *some* of its parts, if those parts are not determined by its substance or essence—hence, the whole determines *some* of its parts. (4) The order and configuration of the parts of *totum essentiale* absolutely matter *a priori*, but in order for such a whole to exist, no change in structure or configuration is possible. The order and configuration of the parts of *totum integrale* absolutely matter *a priori* for only those parts that are determined by its essence.

2.2. 'Modern' Mereology

2.2.1. Leibniz

Leibniz's mereology is central to his entire philosophy.[105] In some ways it is connected to Raymond Lull's system of 'combinatorics', a medieval form of inverse reduction where every whole or compound object can be produced by combining the smallest and simplest parts. In the *Dissertatio* he describes the *ars combinatorial* as the doctrine of the whole and its parts, with the parts themselves being smaller wholes. Burkhardt and Degen refer to the various combinations of the parts of a given set as Leibnizian "permutations." (1990, 6)

Like Aristotle, Leibniz recognises both the "one" and the "heap" and therefore also their corresponding kinds of part-whole relations. Just as in Aristotle, the parts of a heap, the first kind of whole, can exist without the whole, where the whole then is contingent, accidental and aggregate. These aggregates do not have an intrinsic structure and are simply identical with the sum of their parts. With the "one" complete version of the whole (the second type for Leibniz) where an internal structure is present, the whole cannot survive separation from one of its parts. As such, "a man may not survive the loss of his heart or brain." (Burkhardt and Degen, 7) For Leibniz, god[106] is of this second kind, which is also interchangeable with the essential whole (or *totum essentiale*).[107] However there also exists for Leibniz a third kind of whole, which can be understood as a mixed whole composed of both dependent and independent parts. This means that some of the parts are independent of the whole, so some parts therefore are not determined by the whole (*Principes*, GP VI 601, 2).This can be called mixed, but it clearly converges on the use of *integral whole* that we found with the Scholastics.[108]

[105] And thus this short treatment here could not hope to do it full justice, but this is the case with all of the authors herein.
[106] Cf. GP (*Philosophical Correspondence* or *Die philosophischen Schriften*, 1663–85). II 459: Leibniz's division of substances into substantias simplices (*Deus, Angelus, anima*) and substantiate. These are, namely, *substantiatum in unum per se seu substantiam compositam* (*homo*, man), and substantiatum unum per accidens seu aggregatum (heaps). Simple substances have no parts, composed substances are "true" wholes and have parts.
[107] Essential human properties for Leibniz include "being reasonable" (*Characteristica Geometrica* V §25, 151). Furthermore, all geometrical shapes and figures are essential wholes.
[108] In this case, I believe integral is used because it is "integrating" both the essential and aggregate parts, and not because of the Aristotelian notion of inseparability as it has been associated, contemporaneously as well in medieval scholarship, with *integral*.

Leibniz's monadological system is based on the definition of a simple substance (monad) as an entity without parts, the proof of which being that:

> There must be simple substances because there are compound substances; for the compound is nothing else than a collection or aggregatum of simple substances. (*Monadology* §2 VI 607)

His assumption was that there is an infinite number of monads, each with an infinite number of internal states, possibly including perceptions, apperceptions, and appetitions.[109] If all wholes (complexes, aggregates, assemblages, but also essential wholes) are composed of monads, which are infinite, it follows then that "*motis omnibus partibus movetur totum*."[110] There is an ascension of "expansive mereological hereditariness" which seems to indicate that the parts in some sense determine the nature of the whole (Burkhardt and Degen, 1990, 10). This follows from:

> Accurately speaking, matter is not composed of constitutive unities but results from them [*sed ex iis resultat*]. (GP II 268)

The notion that matter arises from the combination of monads suggests that any mereology leads to infinite regression (something that Aristotle deliberately tried to avoid[111]) which Leibniz admits here in this famous passage:

> It follows from the very fact that a mathematical body cannot be analysed into primary constituents that it is also not real but something mental and designates nothing but the possibility of parts, not something actual. A mathematical line, namely, is in this respect like arithmetical unity; in both cases the parts are only possible and completely indefinite. A line is no more the aggregate of the lines into which it can be cut than unity is the aggregate of the fractions into which it can be split up. (GP II 268/L 536) [112]

---

[109] According to Hartz (1992, 512), there are four kinds of objects of perception in Leibniz: illusions (trivial dreams or impressions), aggregates (heaps), corporeal substances (organized groups of monads into organic bodies), substantially-bonded corporeal substances (perfectly continuous).
[110] GM (*Mathematical Correspondence*, 1672–96) V 11
[111] This is demonstrated in the same passage above from *Metaphysics* Z.17, 1041b11-33
[112] In Garber and Rauzy's article, the authors claim that the final phase of Leibniz's position on metaphysics can be summed up as follows: "it is quite enough to have monads as grounding unities and the diverse combinations of these relations as a whole. It is enough for the laws of physics. Enough for the wealth and variety of the world both as it is and as it appears. Enough too if the world really is the intention of God. And lastly, it is enough to account for our pragmatic procedures, in which science abounds." (2004, 38)

The notion of *sed ex iis resultat* (GP II 268) is an interesting twist, especially when paired with the regression problem. Whether or not we subscribe to monads, we are obliged to understand that, without perfectly indivisible, bottomed-out units of analysis, modern physics presently has no reason to suspect that the world is not infinitely divisible. No strings have been observed or detected, and we still have only fragments of understanding of the so-called "quantum foam." [113] Putting aside the essentialist, bottomed-out monads though, we still have *upward* causal constitution, so-called (partial) downward necessitation, and at least partial explanation. Leibniz, as well as the Atomists, would share the same direction to look from: up. Though, as we will discuss in subsequent chapters, we may also need to not only look up into higher levels of analysis but also through time or time-slices. The resulting higher agency or higher *unit* of analysis that we employ on higher *levels* of analysis may be dynamic and therefore non-deterministic.

Several authors on Leibniz have claimed that he may not be a mereological essentialist, despite his claim of the existence of monads. If we construe 'essentialist' as the positing of essential wholes, then he is one. If we construe 'essentialist' as the positing that all objects of perception have an essence, he is not one. The undeniable inclusion of the mixed or integral whole into his mereology prevents us from associating him entirely with Aristotelian essentialism. In §*W* of his *Principes*, he remarks that animals change their (proper) parts, such that change does not involve a change of identity as they are 'impérissables'. Animals are examples of integral wholes:

> Souls do not spontaneously emerge from nothing, and neither do the animals. Animals, however, only develop... they change, with only parts coming and going, which happens gradually and in small bits, but unknowingly and continually according to health.
> (*Principes*, GP VI 601, 2)

This description of an animal as an integral whole is largely what modern science tells us is a description of all animals, including humans. Thus, when we look beyond

[113] The foam—lacking direct observation—has been purported to consist of a frenzy of activity where matter and anti-matter arise (out of the foam itself) and collide, causing immediate mutual destruction. It is from these recent experiments that we might be bold enough to challenge the "ex nihil, nihil fit" claims—and reduce some of the pressure on the big bang as *prima causa*. See Lawrence Krauss (2012) for more on this.

Leibniz' monadological and theological dimensions, we might appreciate his stunning insight into a basic premise of how one might say contemporary mereology should be constructed. Broadly there are four tenets to Leibnizian mereology. (1) All wholes (aggregate, essential or integral) and their parts are composed of the same substance (monads). (2) Following from (1), the whole and its parts belong to the same ontological category. (3) The essential whole is determined by its parts, so the essential whole changes if its parts change. By definition, the integral whole changes if essential parts change. (4) Following from (3), the order and structural configuration of the parts are both irrelevant for essential wholes, whilst they *may* be relevant for integral wholes.

2.2.2. Husserl [114] [115]

Husserl's third section of *Logical Investigations* (LI), entitled "On the Theory of Wholes and Parts", essentially inaugurates the discipline of mereology and formal ontology (from which Leśniewski et al will follow—see note 111). The author himself expresses in the 1913 revision that 'LI3' offers the proper way into his thought, and neglect of this document would have significant impact in the comprehension of the rest of his work. He is one of the key figures in terms of stressing a mereology as a *necessary* component of any philosophical system, and probably the first to create 'merenomies'—or laws of parthood. The specific intention of LI3 is to clarify the relations holding between parts, and the *a priori* possibilities inherent in part-whole relations in general. Connecting his mereology to his phenomenology, though critical in Husserlian studies, is outside the scope of this work. I will however consider how his mereology treats part-whole relations and wholeness "in advance of all empirical

[114] In his *Lectures on Metaphysics,* Husserl's mentor Franz Brentano established a three-fold transcendentalphilosophie: (1) First is ontology, which has three objects: a) the analysis of the senses of being, and 'being' in the proper sense; b) the parts of real being; and c) its causes or principles. (2) Second is theology, which is concerned with the first cause of being and whole of reality. (3) Third is cosmology, which is concerned with the world in its unity and multiplicity, its whole as the end of its parts, and its end as the end of history. In Lectures his mereology distinguishes taxonomically between three kinds of parts: physical (parts of bodies), logical (parts of definitions, can be conceptual or distinctual) and metaphysical (parts above the physical). This is all based on *Lectures on Metaphysics* [(1867) Brentano-Nachlasch Ms. M 96] and adapted from Tymieniecka, A-T (2002), 17. It would seem that Husserl in many ways was clearly influenced by this part-whole framing.

[115] Husserl's *Logical Investigations* has been called "the single most important contribution to realist Aristotelian ontology in the modern period." (Smith, B and K Mulligan, 1982)

instances". [Dermot Moran, ed. in Husserl, 1970 (2001)] This means that part-whole relations have an *a priori* foundation in the "Idea of an object" (LI3 §1), or our reasons for considering simple and complex objects as wholes—that which seems to be integrated and individuated. This also means that Husserl's mereology, like the wider phenomenological programme, does not consider extensionality (i.e. spatiotemporal, real-world objects) and intensionality (i.e. mental or representational objects) as dichotomous.[116]

Every object is or can be a part (*Teil*), assuming it can be distinguished in another object; so, the question is in what manner and under what conditions an object can be one. Husserl distinguishes between essential parts (non-independent, dependent, *unselbstständig*) and non-essential parts (independent, *selbstständig*) via necessity, which are determined by whether or not they can be pulled apart from the whole in order for the whole to remain unified. Moreover, the question is also whether the parts can stand on their own—meaning, the consequences are two-fold: can the whole remain intact after separation of a given part, and can the part stand on its own after separation? This distinction is a return to the ontological notion of dependency, and relation of foundation.

Pieces (*Stücke)* are types of parts that can be independently presented, like segments from an orange; these can (physically, spatiotemporally) stand apart from one another and from the whole. I physically can remove a segment of orange from an orange. Other parts cannot be separated (inseparable, *untrennbar)*, such as regions of colour or movement from a moving body. I physically or mentally cannot remove the orange-ness from the orange without the orange ceasing to be as such. These abstract parts are called 'moments', and a key component of Husserlian phenomenology. So, the handle as well as the colour and shape of a bicycle (with the bicycle being the whole) are all parts of the bicycle. For our purposes, it is difficult to justify equating moments with parts, because we have been until now considering parts that are necessarily

[116] This will be revisited in Chapter IV, vis-à-vis weak object individuation. The unorthodox spelling of *intensionality* and *intensional*, found in recent mereological summary papers (e.g. Libardi, 1994), is meant to reciprocate with *extensionality* and *extensional*: the Latin root *–tens* is presumably used to indicate the direction of 'pulling' of the object in question. Is the object pulled to the mind from (external) observation, or is it pulled in by and (internally) confined to the mind? This is how I understand these terms. Evidently there are vague dualistic parallels with the Cartesian *res extensa* and *res cogitans*.

spatiotemporal (or real-world, extensional objects). Yet at the same time, it would probably not be appropriate to equate Husserlian moments with properties, because properties seemingly cannot be 'contents'.

Barry Smith points out that Husserl is unique in that he is concerned with horizontal parts and not just standard vertical extensionality (e.g. collection and division). These horizontal relations between co-existing parts serve to "give unity or integrity to the wholes in question." In unified wholes, certain parts stand to each other in formal relations of necessary dependence (which is sometimes, but not always, necessary interdependence). Smith continues:

> Such parts, for example the individual instances of hue, saturation and brightness involved in a given instance of colour, cannot, as a matter of necessity, exist, except in association with their complementary parts in a whole of the given type. There is a huge variety of such lateral dependence relations, giving rise to a correspondingly huge variety of different types of whole which more standard approaches of extensional mereology are unable to distinguish. (1998, 20-21)

The notion of phenomenal fusion or *concatenation* is also a key characteristic of Husserl's mereology. This is the relation which holds between two adjacent parts of an extended totality when there is "no qualitative discontinuity between the two." (Smith, 1998, 22) Adjacent squares on a chess-board array, which could be considered dependent parts or moments, are not fused together in this sense. In contrast:

> If there is a band of colour that is subject to a gradual transition from red through orange to yellow, then each region of this band is fused with its immediately adjacent regions and each region is considered dependent. (23)

For example, the parts of a uniform or continuously shaded white surface are independent, but not separated. Such content is fused because they form an 'undifferentiated whole' in the sense that the moments of the one pass 'continuously' into corresponding moments of the other. The acknowledgment of this blurring is in retrospect a double-edged sword: whilst he possesses the insight to see how properties are only roughly approximated (i.e. the inability to differentiate between 'two regions'), he also chooses (perhaps unnecessarily) to consider the continuous surface as unified—in spite of the vagueness. If one assumes, as I do, that the detection of colour gradients alone does not denote integration or metaphysical unity, what ‘rests under’ the object’s matter and force (extensional) to account for its essence? Or, is a surface a whole simply because its contents have a founding relation? But all these questions are dependent on an *a priori* foundation in the “Idea of an object”. Hence, Husserl seems to be able to help us understand the nature of wholes that have moments as contents only if we agree that what we had been calling properties can also be contents.

Aristotelian parts depended on the whole, and thus the whole had ontological priority. The essence determined the whole, and the whole determined the parts. Husserl’s approach to the whole is similar, but he does not pronounce essence *per se* to be real in his mereology. What is real is the founding relation (*Unselbstständigkeit*):

> By a whole we understand a range of contents which are all covered by a single foundation without the help of further contents. The contents of such a range we call its parts. Talk of the singleness of the foundation implies that every content is foundationally connected, whether directly or indirectly, with every content.
> (1970, 475)

The content which is covered by a single foundation or founding relation without the help of further contents (or supplements) is that which is unified. This is his chief sense of *integration*. The nature of this foundation, and the status of its parts as dependent and independent, inform the object’s quiddity and enable *individuation*. His six mereonomies, his rules of parthood, and therefore his means for justifying individuation, are as follows:

> Proposition 1: If an A as such requires to be founded on an M, every whole having an A, but not an M, as a part, requires a similar foundation. (LI3 §14)

This means that A (the part or piece) depends on M (the whole) for A (qua part) to exist. M instantiates A, and if A exists, but not in a given M, A must be instantiated by another in the class of M.

> Proposition 2: A whole which includes a non-independent ‘moment’, without including, as its part, the supplement which that ‘moment’ demands, is likewise non-independent, and is so relatively to every

> superordinate independent whole in which that non-independent 'moment' is contained.

A whole M contains a dependent part (moment) A, but that part can only be instantiated by a different whole W which is an independent part of M. In that case, M is a dependent part of W.

> Proposition 3: If W is an independent part of (and so also relatively to) F, then every independent part w of W also is an independent part of F.

> Proposition 4: If C is a non-independent part of a whole W, it is also a non-independent part of every other whole of which W is a part.

These are the transitivity rules. Prop. 3: If w is part of W, and W is part of F, then w is also part of F. Prop. 4: If C depends on W, C also depends on anything that W depends on.

> Proposition 5: A relatively non-independent object also is absolutely non-independent, whereas a relatively independent object may be non-independent in an absolute sense.

If, relative to a proposition, A is considered as a partially-dependent part of W, A must be (necessarily) absolutely dependent. However, if, relative to a proposition, A is considered as a partially-independent part of W, it is only *possible* (and not necessary) that A is absolutely dependent.

> Proposition 6: If A and B are independent parts of some whole W, they are also independent relatively to one another. For if A required supplementation by B, or any part of B, there would be, in the range of parts determined by W, certain parts (those of B) in which A would be founded. A would therefore not be independent relatively to its whole W.

A and B, both being independent parts of whole W, are both founded by (or having founding relations via) W because they are still parts of W—and are thus prevented from needing supplementation from any other whole. If A or B needed supplementation from a non-W whole, they would be dependent parts of that non-W whole.

Applying our themes, the results vary, from straightforward to tenuous. (1) On the status of the whole's unity, Husserl says that wholes are themselves based on an *a priori* foundation in the "Idea of an object", or what an object can or should be. From there a founding relation (*Unselbstständigkeit*) can be identified that covers and binds a range of contents—themselves being parts, pieces (independent parts), or moments (dependent parts). The founding relation makes the object a whole, and it is what determines whether an object is part of the whole and whether the part is a piece or moment. (2) On the ontological categorisation of both parts and whole, the author holds that distinguishing physical and non-physical ontologies is unnecessary, because all objects are integrated and individuated based on such a founding relation within an "Idea of an object". So, as any ontological categorisation would be crude or irrelevant, one may not need to express the fact that the parts and the whole belong to the same ontological category in virtue of there being no declaration of multiple ontological categories (as opposed to Aristotle). This is the case even if the "Idea of an object" provides for a whole with empirical *and* intensional contents (or moments dependent on pieces or physical parts). (3) On the dependency relation between parts and whole, the whole's dependent contents are determined *a priori* by the nature of the founding relation. Independent contents are parts of the whole, but the whole can survive their loss (the founding relation is still intact). (4) On the relevance of order and structure of parts to whole, because the founding relation determines the order and structure of parts, as well as the categorisation of parts as dependent or independent, the order and structure matter *a priori*.

In sum, Husserl matters for two reasons: (1) he shows, like the essentialists we have covered, that something must *account for* integration, that eventually justifies our individuation; and (2) he demonstrates that integration and individuation, regardless of their means for justification, are dependent on preconceptions of *object* or unit of analysis. He is unique in that the direction of determination—macro to micro, higher-level to lower level, or vice-versa (which Chapter III addresses)—is undeclared, because the founding relation can be known *aprioristically*—and does not need to be contrasted with its material contents.

Mereological questions, as they flow into the metaphysics of emergence, seem to necessarily conflate questions of ontology and epistemology. And, as notions of *whole* are considered, we find our mereological options becoming increasingly more dependent on major metaphysical assumptions. The next group is no different.

2.2.3. The ontological reductionist's whole [117]

First, a few comments are in order to frame what is a contrast from our previous authors. As I discussed in Chapter I, the confidence of late-stage Enlightenment positivism and assumptions of linear scientific progress has never fully dissipated, as demonstrated even in the 20th century: *what there is* can probably be known completely—and numbers and logic will guide the way. These individuals and their successors were wielding the revealing power: nature's laws were to be uncovered, and humanity had the tools to reach out, take them, and use them to our advantage. The whole, as the singular unit of analysis, but also individuated, static, and isolated, and eventually fully reducible to an ultimate bottom level with a full grasp of fundamental universal physical laws—can be known and studied independently of its environment, without perspective, under the safe confines of scientific and/or logical scrutiny.[118] Some of the authors who follow are notoriously sceptical of metaphysics as a discipline, despite the fact that their contributions to philosophy are often based on bold metaphysical assumptions. I have decided to treat this section without all of the same themes as our previous authors, as much of what follows cannot be considered classical mereology. My intention here is to try and understand the platform of modern reductionism with respect to objects and wholes.

[117] Beyond the works mentioned above, the majority of mereological scholarship comes chronologically after Stanislaw Leśniewski's Foundations of a General Theory of Manifolds (1916, in Polish), which itself is, in addition to being the source of the term mereology, the starting point for an axiomatised study of mereonomic relata. English biologist J. H. Woodger's Axiomatic Method in Biology (1937) contained a simplified, English-language version of Leśniewski's theory in an appendix written by Leśniewski student Alfred Tarski. Henry Leonard and Nelson Goodman's article "The Calculus of Individuals" (1940) in the Journal of Symbolic Logic is the first full, published treatment of Leśniewski's theory of manifolds, and together with Whitehead, one can say that formal mereology begins to develop after this time. There is some uncertainty whether Whitehead's theory of events in *Enquiry Concerning the Principles of Natural Knowledge* (1919) may have been formulated prior to Leśniewski's, but nevertheless they remain both important works in CEM. An important difference between mereology and set theory is that mereology is often thought to be conceptually all right if it contains no "atoms": if you can keep breaking things into parts forever. In set theory, unit sets are "atoms" which have no (non-empty) proper parts; many people consider set theory useless or incoherent if sets do not eventually "bottom out" into individuals.

[118] Of course, these latter tendencies have been fading in many circles, and through, for example, complexity studies, we can see the admissions—or seeds of admissions—that what had been considered closed and static may in fact be open and dynamic.

2.2.3.1. Set Theory

A set is a collection of any kind of object, and can be used to collect a wide range of *mathematical* objects. Set theory comes from Cantor's 1874 paper "On a Characteristic Property of All Real Algebraic Numbers", but the basic organising and collecting principles come from Zeno and his followers in an attempt to 'grasp' infinity. We are all members of the set of all humans: that is to say, we either are (1) or are not (0) human. This is standard set theory, a mathematical device with obviously limited human and social application. Applied scientists, social scientists, and set theorists reaching outside the realm of mathematics were faced with problems as a result, and so in the 1960s was introduced *fuzzy set theory*: sets that have members with degrees of membership where a "membership function" is based on a "degree of truth". (Klir, 11) The theorist will have no difficulty contesting the authenticity and completeness of sets that involve anything remotely approaching the individual human level of analysis. The obvious point is that science has for many the dream of confirming that their purported sets were really sets: that all members of that set were undeniably included. This feature is also shared by many analytic philosophers, where trust in the realism of numbers goes largely unquestioned.

Modern mereology is often contrasted with set theory because set theory necessarily has bottomed-out units. Even the axiomatised calculus of Leśniewski was still subject to ontological dependence, a subject Peter Simons (1987) dedicates an entire chapter to in formal predicate logic. The humble and honest position with regards to sets becomes:

- *(1) there must be something outside the set on which at least some members depend on or are affected by in some way (i.e. ontological and condition dependence)*
- *(2) it is conceivable that not all members of the set are in the set (i.e. imperfect observation)*
- *(3) it is conceivable that membership in the set is not reducible to a singular, binary proposition (i.e. incomplete knowledge of properties)*

If the theorist concedes these points, there can be no real sets—at least in the pure, original sense of *set*.[119] If the theorist does not concede these points, they must demonstrate:

- *(A) that the members of the set are ontologically independent*
- *(B) exhaustion of all investigation into what must therefore be a closed system*
- *(C) they have employed axiomatic binary relations using real numbers*

(A) cannot be proven, because knowledge of ontological independence cannot be verified—and claiming so is obviously absurd. (B) is also meaningless because the only closed system is the whole of the cosmos, which cannot be defined nor have its contents known—let alone be the subject of an exhaustive investigation. (C) is standard mathematical procedure, despite the lack of observation of real numbers and despite the inability of human language to be reduced to propositions, and certainly propositions in binary terms. Strangely enough, all three of these are taken for granted in mainstream physics and mathematics, and have seemingly affected certain streams of analytic philosophy for many decades. It seems to me that these are the kinds of deeply-engrained assumptions that are the initial stimulus for opposing emergentism—and not the obvious manifestations of supersummativity (e.g. substance dualism, vitalism, essence, etc.).

The notion that a set is a whole, though, is, fortunately, not as common, and prevents us from laying our themes over it. The reason for the hesitation appears to be, in addition to the obvious issues with fuzzy sets and their definition of membership, the notion that a set does not contain more than its members. It is exclusively a collection: a heap, without the need to call it a heap. It may be a heap of the collection *C*, but it is not necessarily considered a unity—because it is only used in formulation and no ontological claims are necessarily embedded within them. It is the net around the members, which allows us to see the members together. Resisting the urge to call a set a whole is a positive, though hardly a comforting philosophical position. The primary reason for including this section is to speculate into possible motivations for

[119] 'No real sets' does not prevent application of parsimonious sets: instrumentality in mathematics is what designs buildings and standard models, not unfalsifiable claims of real or immutable entities, axioms, and laws.

neat, tidy, ontological reduction, and how mathematics (and all the disciplines that use its tools) might play a key role in hasty integration and individuation.

2.2.3.2. Analytic reductionist wholes

Building on the logical atomistic foundation of Bertrand Russell and others, prototypical reductionists Rescher and Oppenheim in their notable 1955 article "Logical analysis of gestalt concepts" present three features of wholes:

> (1) The whole must possess some attribute in virtue of its status as a whole-an attribute peculiar to it, and characteristic of it as a whole.
>
> (2) The parts of the whole must stand in some special and characteristic relation of dependence with one another; they must satisfy some special condition in virtue of their status as parts of a whole.
>
> (3) The whole must possess some kind of structure, in virtue of which certain specifically structural characteristics pertain to it.
> (1955, 90)

In this essay, the authors admit that they are identifying these features in the context of gestalt psychology—within which there is enormous variation—and it would be outside the scope of this present work to give an adequate treatment of such here. However, the above features do require a statement on ontology, because we need to know if the wholes the *gestaltists* have in mind are primarily a matter of perception or are mind-independent entities. Broadly speaking, the gestalt tendency is to assume that the human mind (brain) is a self-organising system—a whole—with system-level properties. The existence of system-level properties must be assumed for the gestaltists and for any holist—including nominal holists. Furthermore, Rescher and Oppenheim explicitly avoid choosing a definition of a whole as their intentions in the paper appear to be historiographical more than anything else.

I will make some initial comments on these features. In (1), for a whole to be a whole, the whole must have a "peculiar" system-level feature. Immediately, emergentists are led to ask about this peculiarity and its relation to irreducibility and unpredictability, as discussed in Chapter I. In (2), the chicken-or-egg scenario is played out: do we have parts because we have a whole? Or, do we have a whole because we can identify an interrelation between similar parts in the context of an aggregation or alignment?[120] In (3), a whole's structure apparently is derived from the structural properties which exist only in virtue of it being a whole. The confused reader is not alone, as (3) highlights the very problem of talking about wholes: even the use of the word "whole" could imply some feature that is more than (or beyond, or transcendent upon) the parts, even if that something is intensional or abstract.

Perhaps the staunchest opponent of emergentism is Nagel, a pure ontological reductionist, logical positivist, and christener of analytic philosophy. In his heavily-cited 1952 article "Wholes, sums and organic unities", he identifies eight senses of 'whole' in an attempt to be precise about the topics in the title (18-19). I will go through them *seriatim* with commentary for each.

> 1. The word "whole" is used to refer to something with a spatial extension, and anything is then called a "part" of such a whole which is spatially included in it. However, there are several special senses of "whole" and "part" which fall under this head. In the first place, they may refer to specifically spatial properties, so that the whole is then some length, area, or volume which contains as parts lengths, areas, or volumes. In this sense, neither wholes nor parts need be spatially continuous-thus, the United States and its territorial possessions are not a spatially continuous whole, which contains as one of its spatial parts the desert regions which are also not spatially continuous. In the second place, "whole" may refer to a non-spatial property or state of a spatially extended thing, and "part" designates an identical property of some spatial part of the thing. Thus, the electric charge on a body is said to have for its parts the electric charges on spatial parts of the body. In the third place, though sometimes only such spatial properties are counted as parts of a spatial whole which have the same spatial dimensions as the latter, at other times the usage is more liberal. Thus, the surface of a sphere is frequently said to be a part of the sphere, even if on other occasions only volumes in the sphere's interior are so designated.

Especially unlike Husserl, wholes and their parts for Nagel are spatially extended, but not necessarily spatially continuous. This begs at least two questions. A) How can a whole be physically extended, but spatially discontinuous? His example is that the USA is spatially extended, but has parts that are disconnected (referring to outlying islands, etc.). The logical positivist may be overlooking the fact that his whole

[120] Such a dilemma has pitted Parmenides against Democritus, Spinoza against Leibniz, and Hegel against Russell, though it goes far beyond the monism-dualism-pluralism debate. As we will see in the next section, for Aristotle, the *substance* or *form* is what binds the parts, thus creating the whole.

ascription presupposes the definition of the USA as including those islands. The USA has been called such at many points in history, some of which did not include the same discontinuous parts: as such, the ascription of whole to the USA is predicated upon convention. It is not a matter of atemporal logic that the USA is a whole because it has parts that are both spatially extended and non-contiguous. B) We are assuming spatiotemporality for these wholes, and yet we are content to use a crude measurement of space whilst pretending to be precise. How are we to determine the difference between Hawaii and the Pacific Ocean? The Pacific Ocean is not the United States, but Hawaii is. Yet, there is some discrepancy as to where Hawaii ends and the Pacific Ocean begins—and this is without taking into account, for example, the change in sea levels, or the depth of a landmass below sea level (assuming three-dimensional measurement).[121]

> 2. The word "whole" refers to some temporal period, whose parts are temporal intervals in it. As in the case of spatial wholes and parts, temporal ones need not be continuous.

We can allow that a 'time-whole' be considered as one unit of time measurement. One hour is a whole in the sense that 60 minute-parts aggregate temporally and continuously to make the hour complete and unified. If I measure no more than 59 minutes, I cannot claim to have the unit 'hour'. Our definition of 'hour' is a complete sixty minutes, so one does not have an 'hour-whole' with 59 minutes. As such, discontinuous moments cannot be part of an extensional whole, because the completion and thereby definition of the whole requires an observer to decide when the pieces are all assembled. It is easy to see where the author became confused, though. If I am proctoring an exam with a stopwatch, I can stop the clock at thirty minutes past the hour, take a break, and resume the one-hour exam five minutes later. The exam took one hour, being that the total amount of time the students spent taking the exam was one hour. The exam, apparently like in 1), is a whole; as an event, however, its duration was one hour. Nagel is using wholes, then, as events, which is an extremely tenuous proposition. Events are the descriptions of the behaviour of objects over time, which also must be spatiotemporally extended (often in conjunction

[121] And it would need to be three-dimensional measurement, because of the curvature of the earth's surface. There could be no two dimensional mapping in this case, and we are still left with Leibniz' problem of linear infinite regression relative to measurement precision).

with the event itself). Nagel actually confuses time with events, so we must challenge the claim that temporal wholes and parts can be discontinuous.

But, his continuous temporal wholes are not impeccably conceived of either. If we are to be precise with measuring the duration of an event or 'time-whole', we must recognise how time is kept. A whole second is defined as "the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom…at rest at a temperature of 0 K." [122] One needs not concentrate very hard to discover how relativistic this definition is. Which radiation? How much of it? How do we know if we have measured all or enough of it? If time is measured via the ground state of an obscure alkali metal at zero degrees kelvin, surely we can acknowledge the graininess of—and by extension, the parsimony embedded within—these statistical approximations.

> 3. The word "whole" refers to any class, set, or aggregate of elements, and "part" may then designate either any proper subclass of the initial set or any element in the set. Thus, by a part of the whole consisting of all the books printed in the United States during a given year may be understood either all the novels printed that year, or some particular copy of a novel.

This could conceivably be expressed more clearly. What he perhaps means to say is that proper (and exclusively these) parts of a set are necessarily combined to form the whole. It is true *prima facie*, because it has defined itself as such. A 'set-whole' with proper parts has a predetermined set of properties considered necessary for admission into the set. So, Publisher P issued 500 copies of Book X Version A, and 500 copies of Book X Version B. Inclusion by the set-whole Version A requires only that it is Book X Version A and printed by Publisher P. There can only be 500 in this set, and inclusion into the set does not relate to level of degradation or identity of previous ownership of one of those 500 copies. It can work for anything: I crinkle five pieces of paper from the same notepad into five wildly different shapes and sizes, but I decide that they all belong to my set of crinkled papers. It is not metaphysically significant.

[122] From *le Bureau International des Poids et Mesures*, the international system of weights and measures, and the international standard for time-keeping.

> 4. The word "whole" sometimes refers to a property of an object or process, and "part" to some analogous property which stands to the first in certain specified relations. Thus, a force in physics is commonly said to have for its parts or components other forces into which the first can be analysed according to a familiar rule. Similarly, the physical brightness of a surface illuminated by two sources of light is sometimes said to have for one of its parts the brightness associated with one of the sources. In the present sense of the words, a part is not a spatial part of the whole.

A) Parts-as-objects can have properties, but properties do not have parts, because properties are not objects. We cannot say that mass—the most commonly purported intrinsic property—has parts, nor can we say that *luminescence* or *brightness* has parts. B) Brightness[123] is used incorrectly in the example above, as it (at least contemporaneously) is regarded in optical physics as anachronistic, unnecessary and entirely subjective—a confusion of the terms *radiance* and *luminance*—all of which are measured in a so-called RGB (simulated) colour space. All these concepts are observation-dependent, condition-dependent, and are based on constants like the Stefan-Boltzmann law. The now-obsolete equations produced arithmetic mean measurements expressed in units meant for either radiance or luminance, and thus expressed incorrectly.

C) Properties may be observed as a result of one or multiple events involving objects with parts, but properties must always be ascribed to objects, some of which are involved in—or provide agency within—events. Are ("physical") forces wholes, parts, events, or properties? Parts, which are objects, can be wholes, but events and properties cannot be wholes. Forces, as based on the standard model of particle physics, can be all of these things. Photons (electromagnetic force) could be instrumentally seen as objects in the sense that they can hypothetically be isolated and individuated, with potentially autonomous causal powers. Waves are not objects, because, whilst in principle having temporal boundaries, they do not necessarily have spatial boundaries. Bosons (weak and strong forces) are mathematically-predicted but not empirically-observed or isolated force carriers, existing as the analogue to the material fermions. They, like gravity (and the obstinately-postulated *graviton*), have proposed agency in the forms of atomised carriers or quasi-particles, but really just find themselves in the standard model purely to satisfy the need for elegant

[123] With regard to stellar objects, brightness is quantified as apparent magnitude and absolute magnitude—but still subjective and observation-dependent.

mathematical equations in the ill-understood subatomic world. Photons, as objects, and being one of the two forms that electromagnetism manifests itself as, are perhaps the best understood manifestation of force-as-object, and as a result, perhaps, we have good reason to see them as objects. The others, however, cannot be seen as such, largely because we have no empirical reason to hold them as objects in whole-form or part-form. These others must, for now, be seen as properties of a given field or statistical approximations of conditions within subatomic events. D) Nagel writes that these kinds of wholes (4 above) have parts that are not spatial. It would seem obvious that the parts to such wholes would not be spatial, because the erroneously-construed wholes themselves, as events or properties, are not spatially extended.

> 5. The word "whole" may refer to a pattern of relations between certain specified kinds of objects or events, the pattern being capable of embodiment on various occasions and with various modifications. However, "part" may then designate different things in different contexts. It may refer to any one of the elements which are related in that pattern on some occasion of its embodiment. Thus, if a melody (say "Auld Lang Syne") is such a whole, one of its parts is then the first tone that is sounded when the melody is sung on a particular date. Or it may refer to a class of elements which occupy corresponding positions in the pattern in some specified mode of its embodiment. Thus, one of the parts of the melody will then be the class of first notes when "Auld Lang Syne" is sung in the key of G minor. Or the word "part" may refer to a subordinate pattern in the total one. In this case, a part of the melody will be the pattern of tones that occurs in its first four bars.

What is revealed here is the author's selective consideration of time and particulars in the ascription of wholes. A piece of music with written notes is an arrangement of predetermined symbols with predetermined meanings that allow the performer to produce the sounds provided by the notes as close to *reproduction* as possible. The composer wrote the music on paper once, and this writing was reproduced. Performances are never actually replicated, despite the performers' best efforts. The reasons for this can be environmental, personal, technical, or otherwise. But, once we agree that individual performances are particular performances, and not perfect replications or iterations of a universalised set of protocols, we therefore agree that the only wholes that exist are the particular performances and the particular reproductions of the written protocols. A popular song may remain in collective memory, but there is only incongruency between individual memories. The song, as it exists collectively, is not a whole, and is certainly not an extensional one.

> 6. The word "whole" may refer to a process, one of its parts being another process that is some discriminated phase of the more inclusive one. Thus, the process of swallowing is part of the process of eating.

There are numerous challenges here. A) We again have the problem of events as wholes: events are descriptions of the behaviour of objects within a specified duration. Events cannot be considered as wholes because, unlike wholes, events do not have parts. Events have objects, conditions, and an approximation of duration. Wholes must have parts who all share a common connection to the whole; the currency of whole-talk is part-talk. Those parts might themselves be included in events below the level of the whole, or at or below the level of the parts, but a discussion of wholes is necessarily a discussion of parts (not an event the whole or its parts are included in). The temporal dimension of an event does not exclusively belong to the event, because by definition, an event is something that takes place over time. Time does not cease when the event does, so time is not really part of the event. It is the means by which we measure the duration of said event. As such, time cannot be a part of a given event; in fact, a temporal dimension cannot be a part of anything, as it is an observation-dependent substratum[124] of everything.

B) If we grant swallowing to be a 'process', or more precisely and particularly, an event (or 'process' as 'class of event'), we would need to know which objects are included, their properties, and the duration of the swallowing. Of which objects does swallowing involve, what do they do, and how long does it take? Can we be specific about when swallowing begins and when it ends? If one claims to be specific about this duration, how many decimal places in the calculation of the duration? C) Eating does not necessarily require swallowing, if by 'eat' one means ingestion of food by any means. Does eating include intravenous nutrient ingestion? Who decides which definition of eating? Also, do at least some kinds of swallowing not also require brain activity? By Nagel's logic, should we not say that thinking is part of eating? If brain activity is involved, when does the relevant activity begin, and which synapses are we to include in this alleged event-whole? If swallowing is part of eating, surely gland secretion is a

[124] I would propose that this term, which will be used sporadically in this work, be simply understood from its Latin form meaning "to spread under". I am using it as one given perceptual plane: the vaguely-individuated region of space-time that an observer consciously or unconsciously recognises or believes is perceptible in some way.

part of both. If gland secretion is a part of eating, then perspiration is a part of eating. Obviously, Nagel is not really trying here to present a thorough or formal mereology, but his attempts at a taxonomy of 'whole' are probably incomplete.

> 7. The word "whole" may refer to any concrete object, and "part" to any of its properties. In this sense, the character of being cylindrical in shape or being malleable is a part of a given piece of copper wire.

This does not seem to hold much weight, either. (A) Properties are always context- and observation-dependent, unless we are dealing with intrinsic properties. However, as discussed in this work already, intrinsic properties are system-level properties, and require integration, individuation, and, necessarily for many, category (and, as we will see later, individual objects' persistence).[125] Whilst he may be assuming that a given object meets these criteria, our cautious physicalism cannot make such assumptions.

(B) Malleability of copper wire is entirely unspecific, and the term is used to classify metals to estimate its behaviour at standard Earth-bound conditions. The cylindrical shape of a wire is also approximated, and is resultant of a myriad of forces whilst being subject to change. What may be cylindrical at one moment may not be at the next. Furthermore, how many decimal places do we use to determine the circular perfection of each end of the cylinder? If we say ten decimal places, and our two measurements do not perfectly match, we do not have congruent circles on both ends of the cylinder. By definition, we no longer have a cylinder, and we therefore are no longer able to ascribe 'cylindrical' to the wire.[126]

(C) Objects have properties, and only objects can be wholes. A part must be an object, so properties cannot be parts. Not all objects that have properties are necessarily wholes. Furthermore, allowing properties to be parts provides for some strange results. A wooden chair is a concrete object, at least according to common parlance. One of its properties is its weight at standard Earth-bound conditions, vis-à-vis temperature and pressure, for example. However, if, for example, pressure changes,

[125] In Chapter Three, there is an extensive discussion of dynamic properties, including of mass—recently thought to be perhaps the only intrinsic property.
[126] Unless of course if we really want to do so, or if we change the meaning of cylinder. This, of course, has happened many times in philosophy and in science.

so does its weight. How can the chair have dynamic properties as parts? I may say that a particular chair C has property P at t_1 but not at t_2. Does C exist at t_2? In order for us to claim that C does exist at t_2 without property P, we would also need to say that C is without an infinite number of properties that C had at t_1. A runner is dehydrated after a race, though we would not say that dehydration was a part of him. Properties can be rough, observation-dependent approximations of the conditions of an object and/or its parts. The runner was tired, and the runner's legs were tired, so our property is ascribed to either the whole or a part of the whole—and we are not required to classify the kind of part we are ascribing the property to.

In other words, if properties can be parts, then C has an infinite number of parts, because there are an infinite number of properties—particularly at multiple instantiations (and assuming the non-existence of observation neutrality, a fair assumption in post-Einstein physics). If one disputes the claim that there is an infinite number of possible properties of an object at a given instantiation, which would imply that there is a *finite* number of possible properties, one would presumably need not only observation neutrality, but also an absolute, Laplacean, ontologically-independent intelligence with complete knowledge of all ontological and causal dependence.

And, it gets weirder. If we remove a certain amount of wood from the chair, and then we alter the pressure of the space around the chair, we can actually maintain the weight of the original chair to a finite number of decimal places. Though we can make a claim that the chair is a whole at both t_1 and t_2, we would not be able to claim that the two chairs are the same whole. If the two wholes are qualitatively and quantitatively different, they cannot have the same parts at both t_1 and t_2.

> 8. Finally, the word "whole" is often used to refer to any system whose spatial parts stand to each other in various relations of dynamical dependence. Many of the so-called organic unities appear to be systems of this type. However, in the present sense of "whole" a variety of things are customarily designated as its parts. Thus, a system consisting of a mixture of two gases inside a container is frequently, though not always in the same context, said to have for its parts one or more of the following: its spatially extended constituents, such as the two gases and the container; the properties or states of the system or of its spatial parts, such as the mass of the sys-tem or the specific heats of one of the gases; the processes which the system undergoes in

> reaching or maintaining thermodynamical equilibrium; and the spatial or dynamical organisation to which its spatial parts are subject.

Of the possibilities the author lists as parts, I will comment on only the first group, as the others have already been addressed. He says that the mixture's spatially-extended parts, such as the two gases and the container, may be considered as parts. First, why the sudden interest in spatially-extended objects? Surely this is the primary criteria for parts anyway. However, it is not well-expressed, because a gas (or loose, low-density collection of molecules) in its very nature has no spatial boundaries. Tracking each molecule of either kind of gas would be a formidable task, so how can the author claim that the two gases are individuated and distinguishable? The molecules will disperse stochastically in the container, and their behaviour will not be precisely symmetrical nor perfectly reproducible. This is an example of particulars behaving in a certain way at a certain time under conditions specific to a given observation. The other issue is about level of analysis: if we cannot specify the movement and behaviour of each molecule of both types, why assume that we have two distinct units of analysis? Surely the level of analysis is the molecular level, because that is the lowest possible level where the two gases are distinctive. There cannot be two parts, because the behaviour of the gases in the container cannot only be expressed in terms of either gas' agency. We must say that there are more than two parts of the mixture, because the lack of knowledge of the behaviour of all molecules. We may claim to have an approximate knowledge of what is happening in the container, but we cannot say that we know precisely what each and every molecule is doing or has done. Considering the gaseous mixture as a whole can only be done with approximation, where not all the parts are accounted for, yet roughly confined to a space.

It is actually conceivable that the author in this taxonomy is perhaps intentionally painting a bleak picture of wholes, so that his wider programme may be supported. Of course, the end result of this, whether intentional or not, is that not only are his own positions not strengthened, but they may in fact be weakened. This is due to the selectivity of his analysis, and some of his more absurd claims—even with the reduced scientific knowledge in his own time. Both of these analytic contributions above are trying to grapple with complex systems and the first wave of post-Bergsonian emergentism—either the weakly emergent properties revealed by genuine investigation, or new claims or approaches to supersummativity—in the first few decades of the twentieth century. They are trying to dismiss the notion that a whole might be more than what they want it to be: either a bag of ingredients or a Swiss watch. Emergentism, on the other hand, requires an account of the whole that goes beyond the parts and ingredients—beyond mechanistic reduction.

2.3. What a whole can be

The mereologies above are fundamentally embedded within wider philosophical systems, and in some cases actually provide their skeletal structure. What we have seen is that, whilst units of analysis qua wholes are seemingly necessary for everyday language and investigation, the rules of constructing wholes (or determining what a whole is or can be) largely depend on ontological, epistemological, and broader metaphysical assumptions. We are left facing scattered notions of *whole* with vast differences between them and the systems in which they are found. This alone could explain why claims in the metaphysics of emergence are often confusing and seemingly incomplete. More specifically though, we need to make a choice of what a whole can be and how an object comes to exist qua whole. And, by holding a refined physicalist lens, that choice needs to be empirically sound. I will now summarise our options.

2.3.1. Essential vs. aggregate wholes

Defining the nature of the whole is possible especially when rooted in an axiomatised metaphysics. Plato ('version B'), Aristotle, the Scholastics, and Leibniz can claim the existence of an essential whole because their systems use forms and/or substances (essences) that transcend the elemental or physical components. As shown in Figure 2.3.1 below, it starts with the basic proposition that matter emerges from substance (eternal substance, *prima causa*, god, monads). We can then learn about the world by recognising the substances (or forms) and accidents that stand beneath the matter giving it purpose and meaning. This combination yields an essence, and it can be known objectively. Matter-form compounds exist, and their nature is determined by the essence of the thing. As we begin to recognise these compounds, we construct quiddities directly informed by their essences, enabling us to organise these compounds into categories or topics and types of wholes.

Figure 2.3.1.1: Essential Wholes

SUBSTANCE / ACCIDENT / FORM → MATTER

MATTER + {SUBSTANCE / ACCIDENT / FORM} → ESSENCE

⇓

ESSENCE → ESSENTIAL HYLOMORPHIC COMPOUND

⇓

ESSENTIAL HYLOMORPHIC COMPOUND → QUIDDITY

⇓

QUIDDITY → A PRIORI CATEGORISATION / WHOLE TAXONOMIES

In order to claim that there is something which binds the parts into an integrated, individuated whole, one must (a) have empirical or logical reasons for doing so, or (b) claim the supersummative existence of either a spatiotemporal or non-spatiotemporal form or substance-like entity. Let us consider the options for (b). Souls are not spatial, but they are temporal. Aristotelian substances are unobservable directly, but deducible with (his) logic—as they are with Aquinas, who then uses scripture (revelation) and church teaching to confirm that there is such a binding entity. Leibniz follows Aquinas with souls, but challenges both on compositional direction.

Atomism and standard physicalism cannot recognise non-spatiotemporal, unobservable entities (such as substance or essence), despite having no empirical reason to also assume integration and possibly even individuation. In our refined physicalism, we also cannot accept supernatural entities or forces because they are unobservable directly. If that which binds the whole, such as essence, is unobservable, and is yet a real thing, it would require a new ontological category (something other than matter and force). A refined physicalism has to consider this as possible, but unlikely, and currently lacking in observational evidence—and therefore presently off the table despite the explanatory and organisational power *essence* may have. There would remain the question of essential properties—and not a real essence itself—but essential properties are always *refer* to essence, and we are back to where we started. Essential properties, being distinct from intrinsic properties, could be used instrumentally, to aid in categorisation, but would not need to be labelled as such without a real essence to which one would refer.

The mereologies presented have also discussed wholes (or at least compound objects) that lack essential properties (or at least a binding form). How are we to view heaps and aggregate wholes? A heap of sand has no apparent unity, and we would probably be unjustified in claiming that it is integrated. But, we have still referred to it as a single unit of analysis—it seems to have been individuated. It has many parts, none of which are singularly essential to the heap of sand being a heap of sand. The *Sorites Paradox* identifies the problem of whether or not the *soros* can mean anything greater than a single grain of sand. But another issue is of homogeneity: we might have a pile of 1,000 sand grains with one particle of plastic. Is 'it' still a heap of sand? On what grounds do we refer to such a heap as a heap of sand? It lacks a spatiotemporal, part-binding entity, so it does not have a substance or essence. The chemical composition of sand will vary drastically from place to place, so what is sand? If one were to remove all of the igneous rock particles, but leave all of the sedimentary particles, is it only a pile of rocks? Or was that what the heap was anyway?

Let us look at an example on a larger scale. Through the compilation of vast quantities of observational data, astronomers may make claims about our own solar system. It contains a certain number of planets (according to a definition of 'planet'), it has probably only one planet capable of sustaining life as we know it, which is the same planet on which humanity lives, it has gas and ice giants, it has a medium-sized, middle-aged star, it exists about three quarters of the way out from the interior black hole of the galaxy (which itself is travelling at 552 to 630 kilometres per second), which is 100,000–120,000 light-years in diameter containing 200–400 billion stars, within which our system rotates at 220 kilometres per second.[127] We cannot put 'the solar system' with these properties under a normal microscope. [128] The properties that we have designated it with contain or relate to events, objects and other properties, and in everyday language, 'the solar system' refers simply to the collection having these properties. So, upon what grounds might we consider our own solar system as a singular unit of analysis? Is it a whole? Is it a heap? Does it have an essence? [129]

There are many things that cannot be viewed under a microscope, and yet we refer to them as objects or even wholes. If a sample of my blood is drawn and then examined, we still would not say that we are examining 'me'—this is largely because the information gathered from the blood sample does not exhaustively describe its source. However, we may still say that we are studying something if we are only studying a part of it. When the data is presented together in a cohesive form, we may interpret it collectively, as an aggregate. The properties that we have assigned to the solar system are not exhaustive, particularly when there are observations that have not been made yet that may affect the existing set of properties. The objects, events and properties that this incomplete set of properties includes are neither exhaustive nor discrete and static. So this fragmentation seems to challenge both the essence of the 'the one' and the closure of 'the all'.

---

[127] Many of the calculations are based on physical constants, the most important of which is **c**, the speed of light in a vacuum: 299,792,458 metres per second.

[128] In Hacking's words, we could not spray it, therefore making it *non-real*. But how can the solar system not be real, when it has these properties?

[129] A common response to these questions is firstly that the distant, detached, cosmic nature of such a thing prevents us from even considering it as a philosophical device of any kind. We do not know enough about it, we have not experienced enough, it is too far removed from the constraints of everyday language whilst resting largely if not entirely on mathematical and/or axiomatic assumptions. But, such cynicism is connected to the same kind of anthropocentrism that I would argue plagues most social inquiry. Why place a limit of the scale at which we observe, understand, measure, report, discuss, and speculate? If we prefer to philosophise within the range of 10^2 metres, because things in that range are roughly the size of ourselves, then we have promptly dismissed natural enquiry above and below that range, whether or not there may be relevant interactions involved.

The original meaning of 'heap' is a collection of identical, unordered, unorganised, discrete objects. The solar system appears to have order *of some kind*, but we have only vague notions of what that order might be: rotation, approximate mass, composition, and gravitational arrangement. But, these are usually expressed as variables or ranges, and there are quantities that change at the same time that other quantities change. Even the mass of the Earth fluctuates with every measurement. This fluctuation is affected by both our measuring abilities (i.e. local conditions around the measurer) and seemingly objective fluctuations in the actual mass of the Earth (e.g. ozone depletion, loss of heavy elements). Grasping the solar system conceptually requires multivariate analysis, and even then, it is not an exhaustive analysis. The solar system cannot be a heap according to the original definition. If the solar system were to be considered as a whole, according to the definitions we have been using, as a singular unit of analysis, we would need to claim that any explanation or description excludes objects, events and properties outside of it at the time of observation. Who would claim to have that data? Why would someone claim to have that data? How would they obtain it? More importantly, why would they choose to exclude certain kinds of data?

We may concede the homogeneity of the heap, and allow for discrete heterogeneity. That is to say, the heap is composed of discrete objects of different kinds, and only nothingness exists between these particles.[130] There is no order, no configuration, and no design, for it is a soft-sided bag of sand. The solar system as a heap, then, is an approximate collection of matter and force taking different shapes, sizes, arrangements and configurations. If we scale a map of the galaxy up, we can circle our solar system and make this statement about that which is contained within the circle. From this, we can crudely refer to this circle as 'the solar system'. Are there any other options?

It seems that in a refined physicalism ascribing whole-ness to anything must be a compromise. There must be something that binds the parts together which excuses

[130] 'Nothingness', such as it was for the atomists, is that which contains nothing to be analysed. There is nothing between the particles eligible for scrutiny, let alone having nothing upon which the particles depend—which is what is at stake today. Today, we have very little idea what 'nothing' means. See Krauss (2012) for more on this.

our impulse to individuate it. If we say that the solar system, or a planet, or a human being, or even a bag of sand is an essential whole, we are claiming access to its essence as well as its constituency—motivations for individuation prompted by claims for integration. If we claim that these things are just heaps, then there is no room for internal change, either in essence or constituency—through the assumption that there are no external variables. Yet, we need to work with something. We need to have some individual unit of analysis. 'The all' and 'the one' seemingly cannot co-exist, and yet suffer from at least some of the same philosophical flaws. I believe this is where integral wholes and quiddities play a role, because they may be able to grasp some of the uncertainty that a refined physicalism necessarily reveals.

2.3.2. Integral wholes as a compromise?

The Scholastics largely knew that "the all" and "the one" categorisation was not exhaustive. There had to be a way to accommodate objects that found themselves having characteristics of both, before or after *kolobon*. They wanted to retain the non-spatiotemporal (or in some cases, intensional) features of some wholes to supply the informing quiddity whilst recognising the spatiotemporal, unbound constituents. The *totum integrale* presents an opportunity for the more fleeting spatiotemporal components to at least be acknowledged as a part (or content), and not simply a separable grain of sand.

I believe this is a significant breakthrough. Now the spatiotemporally-extended features of a whole help form our conception of *how wholes are constructed*. We are not simply looking at the object's quiddity as being derived from essence, logic, or scripture, in order to individuate and categorise our wholes. The concept of an integral whole is central to this entire work, because theoretically it is no longer blind to dynamic composition and parsimonious integration that can justify experimentally-necessary individuation. In addition, integration and individuation in face of at least microphysical porosity can incorporate observationally-necessary spatiotemporality for explanation: we can now have in principle more powerful dynamic and diachronic explanations to reflect this porosity.

One needs units of analysis *in order to* analyse, and one needs *things actually existing* in space and time to make observations. The *totum integrale*—though for Scholastics and Leibniz ('mixed') possesses a quiddity derived from essence—contains that which allows us to see some sense of unity despite the inconsistency of its physical, observable constituency. It pushes us closer to justifying our impulses for claiming integration and individuation. When holding or poised to hold observational data, which one aims to use in experimentation, we can have a singular unit of analysis despite this obstinate volatility. The key issue, then, is the grounds on which one can claim the existence of such a unifying form *of some kind*—that which justifies integration and individuation, such as quiddity (in quiddity's sense of convention) or, for some, essence (which transcends convention)—whether or not it can be observed or measured, and whether or not it can be recognised by natural investigation.

Wholes, as we have seen, need quiddities, to justify our integration and individuation—even if we instrumentally or parsimoniously exclude levels or units of analysis, and investigate within a finite duration. How do we know *what* the thing is? How do we capture meaning and value from, and ascribe (potentially strongly emergent) properties to, that which we know from observation to be fundamentally disintegrated, porous, interconnected and interdependent? Whilst this is not a new philosophical question, and not one that is confined to the metaphysics of emergence, it has a necessary place within it, and we are forced to engage with it.

And, this question can be approached in various modes and phases. One of the first, perhaps, is to see how wholes are determined. We now have options for what a whole can be taxonomically. So, what are the lessons about whole composition and agency? I had decided to create four themes to give us some structure on these issues. It seems to me that after dissecting these themes, and after extrapolating the results of our survey, we may be justified in establishing a quasi-dichotomy of sorts to frame the nature of both whole compositional determinacy and the source of its causal efficacy—which we will need to do in order to revisit strong claims in the metaphysics of emergence.

2.3.3. Macrocausation vs. microcausation

The next Chapter will address this apparent dichotomy between macro and micro in full detail, but we are at a point in this debate where the nature of integration, being necessary for justifying our individuation (notwithstanding issues of composition and ontological dependence), must inform discussions about what objects are wholes and how we might claim that certain properties of the whole were strongly emergent. The themes presented in most of the selections above specifically include (in numbers three and four) the means by which the nature of the whole is determined. There are predominantly two questions.

- *1a. Does the whole exist qua whole as a result of it being the accidental integration of random aggregative contents, which is individuated via the whole's system-level properties?*
- *1b. Do the contents exist qua parts as a result of them being integrated into or by the whole, which is individuated via its quiddity?*

However, 1a and 1b are distinctly different from these questions:

- *2a. Does the whole depend on its parts for its existence qua whole?*
- *2b. Do the parts depend on the whole for their existence qua parts?*

Aristotelian and Scholastic mereology (and Plato version 'B') would look favourably and affirmatively upon question 1b. This is because, for them, the quiddity (determined by essence) makes the whole, which determines the parts. Husserl would also not mind this, because he would substitute *essence* with *founding relation*. These are the things that bind the whole, giving it meaning and/or purpose, dictating integration, and prescribing individuation. Atomists and modern ontological reductionists would generally favour 1a because the nature of the whole—its quiddity, in our hijacking of the term—is entirely determined by its physical parts (whatever they may be), which when understood together justify our claims for integration and individuation (The Atomists would say that quiddity construction is dependent on our observation of particle aggregation, itself having a sufficient duration). The first set (1) would be preferred over the second set (2) because the whole's existence qua whole and the parts' existence qua parts are both *a posteriori*. This is to say that from these authors above, we can have a whole (roughly) via one of two ways: the whole determines the parts, or the parts determine the whole. Figure 2.3.2 introduces some of the concepts and concerns to come in this thesis, but as we can see, the mereological directionality has profound implications on object integration, individuation, and, by extension, (potentially emergent) property ascription.

Figure 2.3.2.1: Potential implications of mereological directionality

<table>
<tr><th>Aspect</th><th colspan="2">A: Whole → Parts</th><th colspan="2">B: Parts → Whole</th></tr>
<tr><td>(what is affected by this directionality)</td><td>a priori
(is determined by)</td><td>a posteriori
(informs us of)</td><td>a priori
(is determined by)</td><td>a posteriori
(informs us of)</td></tr>
<tr><td>Integration</td><td>• essence
• founding relation</td><td>• category
• topic</td><td rowspan="2">• essence
• conditional dependence
• sufficient duration of aggregation
• chemical structure</td><td rowspan="2">• conditional dependence
• observation
• construction
• level exclusivity</td></tr>
<tr><td>Individuation</td><td>• essence
• founding relation</td><td>• category
• topic</td></tr>
<tr><td>Quiddity</td><td>• essence</td><td>• definition
• convention
• category
• topic</td><td colspan="2">• definition / logic
• convention
• construction</td></tr>
<tr><td>Compositional determinacy (MCD & mCD)</td><td colspan="2">• top to bottom, top-down
• higher to lower levels
• macro to micro</td><td colspan="2">• bottom to top, bottom-up
• lower to higher level
• micro to macro</td></tr>
<tr><td>Causal efficacy (DC, MCP & mCP)</td><td colspan="2">• unified, higher-level agency
• unique causal powers
• possible downwardly-causal powers</td><td colspan="2">• disintegrated, lower-level agency
• aggregative (additive), reducible causal powers</td></tr>
</table>

Some of the hidden questions might be: was the whole there before we recognised its parts (or vice-versa)? If we did not know the quiddity of an object, which even Aristotle and Aquinas admit is a possibility, how are we to know it is a whole? For the atomists and reductionists, which aggregate do we recognise, and at what time? Are the proper parts within our unit of analysis not subject to change? I will need to conclude the chapter here, with compositional and dependence questions beginning to appear that can be framed by macro versus micro.

2.4. Chapter Summary

Connecting the metaphysics of emergence to mereology is necessary because it provides us with options for justifying the integration and individuation of objects to which we might ascribe strongly emergent or reducible properties—with the most prominent set of options simply concerning what a definition of whole per se could be. We started with four general themes for how authors have treated this question: (1) the status of the whole's unity; (2) the ontological categorisation of both parts and whole; (3) dependency relation between parts and whole; (4) relevance of order and structure of parts to whole. Subsequent or synonymous questions would be: (1.1) Should or can the whole be viewed as one? (2.1) Are the parts and the whole composed of the same things? (3.1) How do the parts relate to the whole, and how does the whole relate to its parts? (3.2) Can identity of one only be known through and after the identity of the other? (4.1) Does it matter if the parts within the whole change? (4.2) Can some be added, subtracted, or rearranged?

In going through the selection of authors, I have been able to identify boundaries to what a whole can be, particularly through the lens of our refined physicalism. This had led me to emphasise the importance of the integral whole as an initial compromise between 'the all' and 'the one' that can accommodate the mereological and microphysical volatility that all spatiotemporal objects are vulnerable to. This gets us in the door, because now we see a glimmer of hope that we might be able to retain a quiddity *of some sort* that could eventually justify integration, individuation, and (potentially emergent) property ascription. Without such a quiddity, in a refined physicalism, individuation would be severely threatened, thus upsetting standard investigation: if there are no *units* of analysis, there can be no analysis.

The narrower conclusion of this chapter, however, is the recognition of the importance of mereological directionality, to be expounded in the next Chapter. The systems and taxonomies we had examined showed a preference for either (a) parts that determine the nature of the whole, or (b) wholes that determine the nature of the parts—as shown in Figure 2.3.2. The ramifications for integration, individuation, and quiddity generation are stunning. What I will do next is connect these outcomes more precisely to claims in the metaphysics of emergence: what does a macro-to-micro approach mean for integration? What does a micro-to-macro approach mean for causal capacity? It seems to me that we now need to know how wholes as potentially integrated and individuated objects can come to be.

Chapter III | How is a whole determined?

I have tried to show that emergence needs wholes qua singular units of analysis, and that, in order to justify integration and individuation, those wholes need quiddities of some sort to be (at least seen as) unified. From the literature, we found that such a quiddity originates from a spectrum of claims that can range from (potentially supersummative) essentialism to nominalism—with or without the acknowledgment of the spatiotemporal. In clarifying our mereological options, we also found that, linked to the quiddity's origin, is its mereological directionality: the direction from which a whole is seen as compositionally determined and causally efficacious. As such, the purpose of this present chapter is to show that that quiddity *of some sort* is typically abstracted by either looking up from the lower levels into the higher or looking down from the higher into the lower. I will dissect and try to demystify what is potentially a quasi-dichotomy as we move along.

3.1. Opening Comments

Being that whole ascription is central to strong claims of emergence (as demonstrated in Chapter I), Chapter II leads to the *a posteriori* assessment that the whole's compositional determinacy in classical mereology can take either a macro-level (MCD) or micro-level (*m*CD) path. MCD, where the whole makes the parts, would necessarily require universals, essences, forms, and a number of other concepts that present the object in holistic and potentially autonomous terms—from which the object's quiddity is derived. As its linguistic opposite, *m*CD, where the parts make the whole, could involve atoms or finite, bottomed-out units, but it could also allow for dynamic and diachronic compositional flexibility.[131] All units of analysis require a quiddity *of some sort*, since there needs to be some understanding of what it is that we are studying. This quiddity could come in the form of essence (MCD) or convention-definition

[131] It leaves the door open for weakly persistent and nominalistic explanations, and, as the empirically safer option, it should be an option which merits our attention.

(*m*CD). Typically though, MCD wholes invoke the non-spatiotemporal to justify a universal quiddity. And, as we can see in Figure 3.1.1, this mereological directionality also has implications for causal powers, depending on the origin of the quiddity. In echoing the concerns from the very beginning of the project, we are once again faced with a necessary muddling of the ontological and epistemological. This trend will continue, and, as I have said before previously, I feel there is no way to avoid it.

Figure 3.1.1: Implications of Mereological Directionality

<table>
<tr><th>Aspect</th><th colspan="2">A: Whole → Parts</th><th colspan="2">B: Parts → Whole</th></tr>
<tr><td>(what is affected by this directionality)</td><td>a priori
(is determined by)</td><td>a posteriori
(informs us of)</td><td>a priori
(is determined by)</td><td>a posteriori
(informs us of)</td></tr>
<tr><td>Integration</td><td>• essence
• founding relation</td><td>• category
• topic</td><td rowspan="2">• essence
• conditional dependence
• sufficient duration of aggregation
• chemical structure</td><td rowspan="2">• conditional dependence
• observation
• construction
• level exclusivity</td></tr>
<tr><td>Individuation</td><td>• essence
• founding relation</td><td>• category
• topic</td></tr>
<tr><td>Quiddity</td><td>• essence</td><td>• definition
• convention
• category
• topic</td><td colspan="2">• definition / logic
• convention
• construction</td></tr>
<tr><td>Compositional determinacy (MCD & mCD)</td><td colspan="2">• top to bottom, top-down
• higher to lower levels
• macro to micro</td><td colspan="2">• bottom to top, bottom-up
• lower to higher level
• micro to macro</td></tr>
<tr><td>Causal efficacy (DC, MCP & mCP)</td><td colspan="2">• unified, higher-level agency
• unique causal powers
• possible downwardly-causal powers</td><td colspan="2">• disintegrated, lower-level agency
• aggregative (additive), reducible causal powers</td></tr>
</table>

The figure above, whilst simplified, shows us that at least most of our mereological options in describing and identifying wholes—to which we might ascribe strongly emergent or resultant properties—depends on how this question is answered. Of course, as we have also said, the directionality of composition (which can directly or indirectly justify integration and individuation) and causal efficacy is determined by the object's quiddity. Towards the end of this Chapter, and into Chapter IV, I will specifically address the relationship between quiddity generation and directionality. For now, let us take a closer look at this seemingly dualistic directionality.

3.1.1. Macro-micro clarification

At the conclusion of Chapter II, I extrapolated from classical extensional mereology a vague and general dichotomy of options for understanding how a whole comes to be. This begins a series of possibly confusing usages of 'macro' and 'micro' terminology. Figure 3.1.2 below tries to distinguish variations of these, as I will henceforth use symbols for each variation.

Figure 3.1.1.1: Macro-micro clarity

Aspect of whole (W)	Symbol	Description in terms of parts and W
Micro-level (*m*C) (lower levels)	*m*CD	Parts determine what W is compositionally
	*m*CP	W's alleged causal powers can be reduced to the powers of its parts
Macro-level (MC) (higher level)	MCD	W determines what the parts are
	MCP	W has autonomous causal powers in the world
	DC	MCP + reflexive

3.1.1.1. Composition (D)

We must consider the compositional determinacy of wholes, hence the suffix and symbol ‘CD’. CD is the means by which we individuate. Hume realised the need to separate and individuate in order to establish causal relations and eventually explanations. We know that investigation needs units of analysis, which are necessarily individuated after at least sufficient integration (some positions assume *de facto* integration). And, I believe this is a point where we will encounter some friction. I have already mentioned how claims in the metaphysics of strong emergence and reduction entangle the ontological and the epistemological, and this compositional issue is certainly a key intersection.

The question of how the whole’s composition comes to be determined finds itself posed after an acknowledgment that something (at least sufficiently) integrated actually exists. Initially, we can set the integration question aside, because we are looking at the implications of this directionality for how there can be an X and a non-X—or how a whole is individuated. Later we will see how individuation (object exclusivity) is ultimately bound up with integration (unity) when we are holding our refined physicalist and empiricist lens. Thus, the first type of question we are asking is along the lines of whether (a) the tree permits the existence of its leaves and branches or (b) the existence of the leaves and branches permits the existence of the tree.

MCD signifies that the whole determines what its parts are. Because for the essentialists there exists a unifying form, with which matter combines to make hylomorphic compounds, individuation begins from the top and goes as far to the bottom as possible or necessary. The whole possess this essence, and it can lose certain parts, but in retaining its essence, it retains its wholeness. This is possible because it is top-down: individuation via wholeness, and wholeness via essence.[132] The symbol *m*CD signifies that the parts determine what the whole is. This might appear on the surface to be clear enough, but let us expound. This can be the *soros*,

[132] Integration is only relevant if an essential part is removed, which would then require the whole to cease existing as such. The example from Chapter II was the removing of the brain or heart from a human body: without these, it ceases to be a human whole, because the essence of ‘human’ demands the existence of ‘heart’ and ‘brain’ qua parts.

or it might also be something else. The only criterion for *m*CD is directionality: higher levels of analysis are informed and constituted by lower levels. The criterion for MCD is more than directionality: it must, in virtue of its individuation via essence, claim a unifying element in addition to the direction of composition. I hope this will become clearer over the next several sections.

3.1.1.2. Causal powers (P)

I am using MCD and *m*CD to identify the rough dichotomy between types of compositional determinacy, and this is distinguished from MCP and *m*CP that signify outward causal powers. MCP signifies the power of the object qua whole to instantiate new properties in the world. A particular tree T, as a tree-whole, provides shade for various creatures that then become relaxed and write interesting stories. As such, so the MCP argument goes, T qua whole has autonomous causal powers. T has provided for properties that would not otherwise have existed had T not existed: by extension, these new properties are ontologically dependent on T. *m*CP, however, signifies the manifestation of perceived novel properties via the aggregation of lower-level causal forces. T's apparent ability to instantiate new properties directly and exclusively results from the causal powers of whatever constitutes T—including other objects, events, properties, and conditions. In physical terms, the relaxation-inducing shade provided by the constructed, aggregated organism—previously referred to as T—is caused by partial blockage of photons by more dense objects spatiotemporally situated between the light source and the brain observing and responding to it (and the variation in temperature and its biochemical effects). One is evidently less romantic than the other.

3.1.2. Compositional determination versus causal capacity

Let us briefly look at why it is important to distinguish between the composition and the causal powers of both the whole and its parts. Much of the classical mereology that had been presented in Chapter II speaks exclusively of whole composition, and not the actions of the whole now that it has been recognised as such. Talk of essential wholes is not necessarily talk of causally-efficacious wholes. And, this is one of the major crises in the metaphysics of emergence. Causal powers of the whole or its parts have been separated from their composition. We, now, are distinguishing them, to specify the difference. It is an entirely different thing to separate them as if they were unrelated. The bulk of the literature in the metaphysics of emergence does not talk about composition (whether MCD or *m*CD) whilst focusing mainly on causal capacity (whether MCP or *m*CP). But, the mereological literature draws predominantly on composition, and not causal capacity. They both need each other, so it is important to show how they are connected. This is not easy, but will hopefully become clearer as we go along.

MCD, again, concerns itself with how the whole is compositionally determined. This means, more precisely, that the unit of analysis is known only because the higher levels have incorporated the lower levels. It is from the perspective of the higher level that we recognise it as a whole—which must have discernible parts and/or lower levels. *m*CD to some degree reverses this: the unit of analysis is known only because the objects on lower levels co-ordinate themselves spatiotemporally. This co-ordination is recognisable on a higher level of analysis, and when a 'configured' macro-object is found in this same higher level of analysis, there becomes a single unit of analysis—the whole. When described in this way, the MCD-*m*CD dichotomy seems to reveal an intersection of epistemological and ontological claims: what we recognise as existing is also interwoven with how we recognise it.

In etiological discussions, the MCP whole is a robust, active unit of analysis. Humans interact with the tree, but the tree interacts with us. We draw and place the blood sample under the microscope, but the sample creates changes in us. Before going too far into this, as I will do in the next few sections, we can already see the problems with claiming that the blood sample creates changes in us. Let us for the sake of discussion grant that the tree is to be vaguely identified as a whole with outwardly causal powers. What properties have been instantiated by what means? The argument might be that the macro-object might be participating along with the human in some sort of "dance of agency”,[133] whereby both interact through feedback loops all connected to the matter and force in, through, around, and between them. Some things that we ‘do’ affect the tree: this might be something as simple as *leaning* on it, where the pressure of the leaning body affects a relatively small group of exterior cells (which then feed back into the allocation of resources). Some things the tree does affect us: the edible fruit of the tree is picked and consumed, nourishing the human and his vital organs. With *m*CP descriptions and explanations, it would be not be the tree qua whole that possesses these powers. *m*CP does not recognise the tree as having its own autonomous causal powers. One recognises the tree as a singular unit of analysis, either actually or parsimoniously, but explains the interactions as above in lower-level terms. The nourishment of the human’s normally-functioning brain relied on the programmed distribution of the vitamins and minerals found via the fruit’s digestion throughout the body in the bloodstream. Furthermore, the entire series of anatomical events was initially caused by the decision—the firing of synapses—to ingest what appeared to him to be a palatable source of vitamins. In other words, in *m*CP terms, there are, quite possibly, an infinite number of correlative causes that stand to be recognised before the claim in this instance of the tree’s MCP.

The micro and macro will battle over questions, in causal and/or compositional fashion. It will be difficult to reconcile them, too. The metaphysics of emergence, on the surface, recognises the uniqueness of the macro-level. And, surely the vision of the human qua cellular automaton is far from tenable. But what kinds of constitutive and causal capacities are we willing to grant emergent, macro-level wholes? Some have gone very far indeed, as we will see in §2 with a unique form of MCD: downward causation. First though, there is the obstinate issue of change over time for both CP and CD.

3.1.3. Initial comments on the persistence of wholes

I will now briefly comment on the important relationship between a whole’s existence and a whole’s persistence, because, as we will see, justifying our integration and

[133] This phrase is from Andy Pickering’s *The Mangle of Practice* (1995).

individuation (either top-down or bottom-up) through a refined physicalist and empirical lens must acknowledge the diachronic and dynamic nature of (at least the observation of) the whole. This section is merely intended to hint at the overlapping nature of the problem, where a few ideas can be gently introduced in the hopes that the broader picture of the thesis might become clearer. As we consider the more precise implications of our various CD and CP options, it may be with some regret that we will find that engaging with persistence, despite it being itself a robust sub-field of contemporary metaphysics, now seems to be necessary. Though this is the primary focus of Chapter IV, the present chapter's analysis of wholes will begin to incorporate the themes of time and persistence and their relation to integration and individuation.

Examples of wholes possessing essences, such as a human person,[134] would necessarily need to exist between multiple observations (e.g. t_1 to t_2). Whilst many mereologies address the question of the addition and subtraction of parts (survivability), most mereologies do not address the issue of *the means by which* the whole in question actually persists across multiple instantiations—though this is largely because is it an *a priori* issue of whether a whole has survived. In other words, from this *a priori* perspective, a whole exists across observations because it is still the same whole with the same parts throughout. Through a refined physicalist and empiricist lens, however, it may not be so easy to assume that all of the parts at t_1 are present at t_2. If the constituency fluctuates, the core notions of integration and individuation are placed in doubt. Thus, we enter into the realm of persistence: by what means does a whole persist, and what criteria exist for the whole at t_1 to be the same as the whole at t_2.

Within the identity/temporal parts domain of contemporary metaphysics, two sides often emerge as being pitted against each other: endurance and perdurance.[135] I will

[134] I phrase it in such a way because some mereologies would not consider a human person an essential whole, but rather as a mixed or integral whole that has an essence.
[135] This is a plain definition of a whole that wholly persists through time, or at least from one instantiation to another. Much of the endurance-perdurance debate originates from David Lewis (1986). For all references to 'endurance' henceforth, I am using a standard three-dimensional view of endurance, such as from J. J. Thompson (1983). For all general conceptions of 'perdurance', I am considering several variations in four-dimensionalism herein, and will not, unless otherwise hinted, disqualify either stage or worm views. A more thorough treatment, where persistence is primary focus, comes in the next chapter.

spend considerable time on perdurance (or at least on four-dimensionalism) and general persistence options in the next Chapter, but I need to address what is at stake now before my treatments of MCP and MCD wholes can develop any further. Very briefly, endurantists claim that a given object wholly persists from t_1 to t_2, where nothing or no part is spared. Perdurantists, on the other hand, hold that all objects have temporal parts—where temporal parts are “parts that exist solely because of the object’s existing at a certain time” (Simons & Melia, 2000, 59-60)—which compose the object at specific time-slices or instantiations. I will in this present Chapter begin to engage with the consequences for wholes that persist in these and other ways.

In classical mereology, wholes are assumed to endure, because they are that of which nothing or no part is missing at any given instantiation, and possibly even that which is something that has not lost nor will lose anything over the course of its existence. The whole and its parts, together in the forming of the continental notion of a continuant, being wholly present throughout their existence, are objects that exist through or in time, with no temporal parts. We have already discussed the prospects of theories that assume objects do not change—or more specifically, to assume that objects cannot have different properties at different instantiations—and they will not be discussed here. But, the usage and understanding of temporal parts are seemingly diverse and/or selective, and because of this, endurance seems to, as it were, endure—especially in an *a priori* or mereological context.

In our discussions of CD and CP, there may be a few predictable alignments. For example, macro-level compositional determinacy and causal efficacy will seem to coincide with endurance, because the whole is assumed to exhibit the same system-level properties from t_1 to t_2. Micro-level CD and CP will be difficult to analyse through this wholly-persisting lens, because of its necessary acknowledgment of the microphysical. In any case, let us look first at the top-down approach and the implications for integration and individuation.

3.2. Downward and macro-level causation

Some strong and ontological forms of emergence recognise irreducibly novel properties featuring downward causation[136] (DC), where the autonomous whole seems to be able to control and affect its own parts. Pure, ontological emergence as well as DC, has been consistently challenged, even by Kim, whose subdued acceptance of a form of MCP will be discussed later. The obvious problem with OE[137] is that it would introduce a new kind of entity into the world composed of things not recognised by physics—an idea currently lacking in any observational evidence. Since we are unlikely to uncover a class or instance of such an entity, we are therefore left to deal with DC and MCP. DC requires the following to be true of a given object qua whole:

- *(1) The extensional whole endures.*
- *(2) The extensional, enduring whole has causal powers that are somehow dependent on, yet autonomous from, its constituent parts.*
- *(3) The extensional, enduring whole has the causal capacity to instantiate new properties in the world as well as within itself.*

How can the whole have properties disassociated from that which it is made of? On the surface, this is perhaps the core of emergence, and it is what we have been addressing since the beginning of this present work. But, these demands of DC are very steep, and are multi-dimensional. A DC whole must endure, which is already a serious issue, but it also must have powers that do not result from the powers of the parts. Furthermore, it must have the ability to change its own constituency. Amongst other issues, there may a problem here of reconciling *autonomy*, or disassociation, with dependence. In any case, all three of these demands will be treated herein, because they connect back to DC's parent concept of MCP. First though, we must treat DC.

---

[136] There is another version—so-called unpredictable downward causation—that a few authors have espoused, but it is an unnecessary complication and will not be dealt with here (as it seems to be a combination of 'explanatory incompressibility' and macrocausation).
[137] The reader is reminded that, in Chapter I, I decided to distinguish between ontological emergence (new class of things with DC, etc.) and strong emergence (DC, MCP, MCD). So references to OE are for the former.

3.2.1. The challenges of downward causation

The most provocative feature of DC, as identified and challenged by Kim, is its reflexivity: the proposition that that which can be called 'whole' at t_1 can exert independent causal influence upon its own constituency also at t_1. Roger Sperry's work on mental and macro-level causation advances this reflexive form of DC, or as he calls it, "direct macro-determinism" (1986, 266). Here are two excerpts:

1. "Recall that a molecule in many respects is the master of its inner atoms and electrons. The latter are hauled and forced about in chemical interactions by the overall configurational properties of the whole molecule. At the same time, if our given molecule is itself part of a single-celled organism such as a paramecium, it in turn is obliged, with all its parts and its partners, to follow along a trail of events in time and space determined largely by the extrinsic overall dynamics of *Paramecium caudatum*. When it comes to brains, remember that the simpler electric, atomic, molecular, and cellular forces and laws, though still present and operating, have been superseded by the configurational forces of higher-level mechanisms. At the top, in the human brain, these include the powers of perception, cognition, reason, judgment, and the like, the operational, causal effects and forces of which are equally or more potent in brain dynamics than are the outclassed inner chemical forces." (1964, 16)
2. The subjective mental phenomena are conceived to influence and govern the flow of nerve impulse traffic by virtue of their encompassing emergent properties. Individual nerve impulses and other excitatory components of a cerebral activity pattern are simply carried along or shunted this way and that by the prevailing overall dynamics of the whole active process (in principle - just as drops of water are carried along by a local eddy in a stream or the way the molecules and atoms of a wheel are carried along when it rolls downhill, regardless of whether the individual molecules and atoms happen to like it or not). (1969, 534)

In these examples, Sperry is using a quasi-essentialist MCD mereology to integrate and individuate the whole: the macro determines and controls its micro, because contained within the whole is a unifying force that could for our purposes be initially called *essence*. In a more contemporary approach, the word might be, as it is for Sperry, *configuration*, because it justifiably suggests, for the paramecium case alone, that the single-celled organism has a discernible structure and programme. A cell must have a structure (i.e. its capsule, cell wall, or membrane) and a programme (DNA), and together these two features might allow us to claim that it possesses some kind of configurational 'force'. This 'force', though, is not a separate entity, nor is it an essence *per se*. One can call it the cell's 'configuration' only because of the arrangement of macromolecules in a certain way as directed by the DNA programme under suitable environmental conditions.[138] If it is not a quiddity-informing essence, then is it an object at all? Can a configuration be a *something*? It seems to me that this configurational force is first and foremost a system-level property, and not an object *per se*. Because it is not an object in and of itself, it cannot be an essence (at least in the same way it would be for the essentialists).

So, should we say that the paramecium case, assuming integration and individuation as the author does, demonstrates supersummativity, where there actually is more than the catalogued sum of the parts? Or, is its perceived configuration actually just the simultaneous alignment of smaller parts who themselves have a common programme? Bedau answers with another molecular case:

> Another very simple example of an alleged emergent phenomenon is the spontaneous formation and subsequent growth and division of tiny structures, such as vesicles. Vesicles are microscopic hollow spherical bodies, formed from a thin membrane, with a watery fluid inside and out. The vesicle membrane forms spontaneously when certain special molecules, called amphiphiles, are mixed in water. Amphiphiles are polar molecules; they have a hydrophilic end that tends to move toward water, and a hydrophobic end that tends to move away from water. So, when amphiphiles are mixed in an aqueous solution, they spontaneously clump together in certain shapes so that the hydrophilic heads are in contact with water, and the hydrophobic tails are not. One such aggregation is a bilayer membrane. These membranes can be millions of amphiphiles wide, but they are only two amphiphiles thick. The hydrophilic ends of the amphiphiles are all facing outside the membrane, toward the water, and the hydrophobic ends are all facing inside the membrane, between the two layers. In this example, the vesicle is a "whole," and the amphiphilic molecules are its "parts." (2011, 92)

This shows that the configurational force of the vesicle at any given time is largely due to the nature of the chemistry of the molecules inside of it. Needless to say, our knowledge of a particular single-celled organism would be far more expansive than that of a particular human brain. Sperry has, along with contemporary neuroscience, no clear idea what a singular configurational force of the brain might actually be in physical terms. He simply advances the argument that the brain must have some configurational force—it must be supersummative—because higher brain functions seem to possess the power to instantiate new properties in the brain's lower-level parts and processes. However, if we cannot give a thorough explanation of the macro-level

---

[138] Molecules, micro (e.g. H_20) or macro (RNA), might be said to have structures, but they cannot be said to have programmes. In the RNA case, the macromolecule *is* the programming agent, but it does not inherit an aggregated programme from its component nucleotides. If an atom is said to have a structure, it certainly does not have one in the same way that molecules have structures. Neither molecules nor atoms have programmes, but they are composed of smaller particles or units of analysis.

brain processes, how can we de-value the microphysical, which apparently have been "outclassed" by the mysterious macro? Clearly there is a difference of *agency*[139] between the brain, the paramecium, and the atom, but one major problem is Sperry's selectivity: he is ascribing a configurational force to the relatively large objects, but not to the small within the same example. If the brain's cells collectively possess a configurational force (in much the same way as the 'emergence' of the shape of the aggregation of amphiphiles), then why can that singular force not be explained in relation to the configurational forces within individual brain cells? If not, then the claim could conceivably be interpreted as meaning that the prestigious macro-level brain processes do not have any relationship with the programmes and structures of brain cells—an apparent absurdity. Kim, who as we will see later makes the same mistake with another case, recognises Sperry's error in the eddy example:

> After all, an eddy is there because the individual water molecules constituting it are swirling around in a circular motion in a certain way; in fact, an eddy is nothing but these water molecules moving in this particular pattern. Take away the water molecules, and you have taken away the eddy: there cannot be a disembodied eddy still swirling around without any water molecules! Thus, reflexive downward causation is combined with upward determination. When each and every molecule in a puddle of water begins to move in an appropriate way, only then will there be an eddy of water. (1999, 27)

Kim dismisses the reflexive form and hints that it cannot be the sole explanation for the eddy and other phenomena, based on the idea that there irrevocably must be some account of *m*CP. As we will see in the first example below, this reliance on the microphysical for Kim will be short-lived. Kim, along with most authors in the metaphysics of emergence, assumes object integration and individuation, at least in the *a priori* sense. This assumption is precisely why I have been compelled to use a refined physicalism, and not a bottomed-out, closed-system one.

3.2.2. Relating DC to MCP

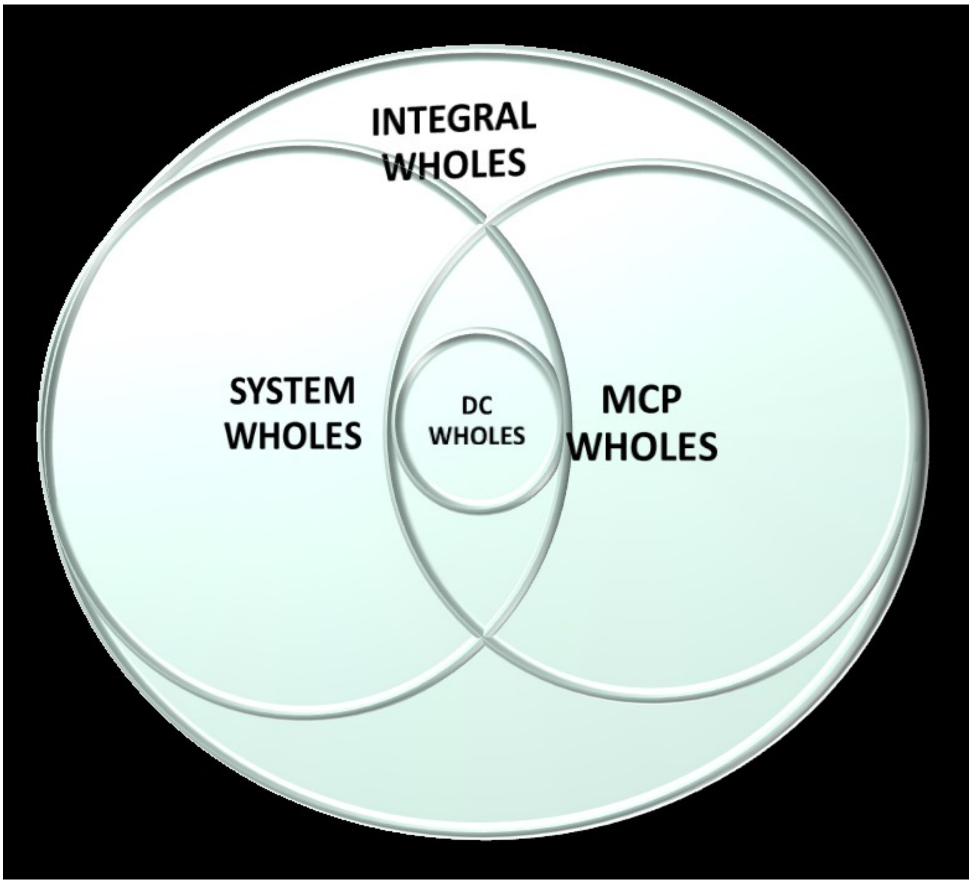
A DC whole at t_1 has the power to affect its own constituency also at t_1, as described in Sperry's autonomous molecules that drag the atoms and electrons along for the

[139] For the moment, let us progress with agency meaning 'causal capacity', or the capacity to instantiate new properties in the world.

ride. Presumably Sperry is also claiming macro-level outwardly causal powers at t_1: the molecule has the ability to affect other things in the world in virtue of it being a whole. All DC wholes must also be MCP wholes, but all MCP wholes are not necessarily DC wholes.[140] The reflexive power of the DC whole requires first the acknowledgment of the higher level as encapsulating the lower levels. Secondly it requires the acknowledgment of this encapsulation as being an autonomous unit of analysis and a whole. If its higher levels can instantiate new properties in the lower levels, it must also be able to instantiate new properties in the world external to the encapsulation. To have inwardly-facing causal powers without outwardly-facing powers is inconceivable. Proving that a molecule M cannot under any circumstances influence the cell C at t_1, which is an already futile task, first of all requires that M is external to C. This is because M's 'dead' state would affect the complex system—the network—that is C. C's components and system-level features would and must be affected by all components of C. Hence, all DC wholes are necessarily MCP wholes (see Figure 3.2.1).

[140] The claims for DC (Sperry, Davidson, Silberstein, others) are far less numerous than those of MCP, which are widespread, in many forms, in many fields, in many different applications.

Figure 3.2.2.1: Strongly-emergent whole venn



For Sperry, a molecule—as an autonomous macro-level whole at t_1—also has the ability to instantiate new properties in the world also at t_1. The evidence for this is simple: he grants no autonomy for the electrons and atoms found in the molecule. As such, he is ascribing holistic properties to the higher level at t_1 and not the lower, and therefore the molecules as they exist within a cell at t_1 have no autonomy in the cell. He ascends in level of analysis to declare that the brain's constituents have been "superseded by the configurational forces of higher-level mechanisms." (1969, 534) This comment comes immediately after declaring the autonomy of the molecule. In other words, he wants to actively promote the higher level as the exclusive agent.

But is it not obvious that a single unit of analysis cannot have autonomy in one sense but not in another? How can the molecule be the master in one sense but the slave in another? It seems to me that any given level of analysis, such as the molecular level, is only the master (autonomous) if it has no dependence on either of its respectively higher or lower levels. Claims in the metaphysics of emergence obviously provide for both directions of dependence. Now, if a strongly emergentist claim considers the molecule as strongly emergent, or having strongly emergent properties, including macro-level autonomy, irreducibility, and disassociation of powers, then anything that the molecule composes could in principle be reduced to these autonomous molecules. This is because the cell (for example) within which the autonomous molecule 'exists', whilst being dependent on the molecule for its existence, it is also subject to its autonomous behaviour. So, the properties of the cell are really just the sum of the properties of the autonomous molecules. In other words, it could be said that being selective about identifying autonomous units of analysis can require ontological reduction from respectively higher-levels of analysis.

If the claim is that his molecule is a DC whole only when discussing his molecule—but not when discussing his molecule in a given context—then he is seeking to isolate the molecule in a vacuum. He would need to do that with all higher levels: the cell must therefore be the master of the molecules only when we are only looking at the cell by, in and of itself. But, as we have already realised, even the proto-cell lives not in isolation. It is, first and foremost dependent on 'external' nutrients in a suitable range of conditions (like what we call 'temperature' and 'pressure'). Strangely enough, he is talking about any molecule anywhere: so, in the context of the cells of an organ, let us say, the molecules feedback into their own constituents, but not to other molecules, and not to the cells. Sperry's account is nonetheless self-contradictory.

The brain for many is used as not only a case for strong emergence in some vague sense, but specifically for emergentist MCP and sometimes even DC. Let us keep it in Sperry's words and ask: how can the brain be the master of the vast range of brain cells, chemicals, molecules, and electrons (to stop there), when the brain relies on, amongst *many* other things, oxygen from the environment? The oxygen does not interact with the brain qua whole: it is filtered through the various systems where the resources are allocated appropriately. The explanation of brain function has never been more clearly looking upward into the higher levels. Like so many others, he is basically arguing that the molecule has DC powers whenever we want it to,[141] as long

[141] This reminds one of Phillip Kitcher's evocative phrase: "a gene is whatever a competent biologist chooses to call a gene." (1992, 131)

as the unit of analysis can be simplified and isolated. Claims for DC can sometimes seem to be rather selective about when to view the unit of analysis in its role as—or its relationship with—a complex system.

The notion that the brain is a reflexive DC whole is entirely preposterous. There is, however, some question as to whether we might call the brain and other complex systems MCP wholes. This is at least initially though perhaps not exclusively because we can now remove the temporal restrictions that had plagued DC, and look at a potential MCP as a dynamic, complex system from t_1 to t_2.

3.3. Cases for macrocausation in MCP form

There is a multitude of potential cases for MCP, with some being what we would call simple and some we would call complex. The primary differences between the two appear to be the relative difference in constitutive heterogeneity (a flower vase compared to a human body) as well as (what we claim to be) the quantity or quality of visible internal and external change over time. My intention here in this section is to show some of the challenges in claiming the *de facto* autonomy of objects qua wholes. Let us start with a simple case for MCP.

3.3.1. Case 1: The Vase

In the case of Kim's vase (V), one can see a kind of 'non-reflexive' MCP, where "an event involving a whole causes other events involving lower-level entities that are not among the whole's constituents":

> The vase on my desk has a mass of 1 kilogram. If it is dropped out the window of my second floor office, it will crash on the paved sidewalk, causing myriads of molecules of all sorts to violently fly away in every which direction. Even before it hits the ground, it will cut a rapid downward swath, causing all sorts of disturbance among the local air molecules. And these effects are surely micro and lower-level in relation to the fall of an object with a mass of 1 kilogram. Note that we cannot think of this case as one in which the "real" causal process occurs at the micro-level, between the micro-constituents of the vase

> and the air molecules, for the simple reason that no micro-constituents of the vase, in fact no proper part, of my celadon vase has a mass of 1 kilogram. There is no question that the vase, in virtue of having this mass, has a set of causal powers that none of its micro-constituents have. The causal powers that this property represents cannot be reduced to the causal powers of micro-constituents of its bearers. Of course, emergentists would not consider mass an emergent property; they would say that the mass of an object is a resultant property, a property that is merely "additive or subtractive". But this simple example suffices to show that there need not be anything strange or incoherent in the idea of downward causation as such - the idea that complex systems,[142] in virtue of their macro-level properties, can cause changes at lower micro-levels. (1999a, 25)

To clarify, Kim here wants to grant autonomous causal powers to V in question. For him, it is a whole, which apparently remains constant in mass and composition from the moment it was dropped (t_1) to its approximate half-way point (t_2) up until the moment it crashes (t_3). However, the description of the event and the 'whole' within it seems to be incomplete, and so the conclusion may be incorrect. Though the example identifies *some* physical properties, conditions, and events, the most important ones seem to have been overlooked.

3.3.1.1. Compositional omissions

The first set of omissions concerns the composition of V. The author has claimed that its mass is 1kg, and apparently he is implying that V remains 1kg from t_1 up until t_3. However, he fails to recognise that the mass of any particular object, though a so-called *intrinsic* property, changes over time. Any calculation of mass of a particular object at any time will be different than the 'same object' at another time, with the only limitation being the number of decimal places. Considering the earth's relatively stable physical conditions, measurements of V at t_1, t_2, and t_3 will easily remain 1kg with standard measuring devices. In reality, the mass of the object from t_1 to t_3 is more likely to be between 0.999999…9 kg and 1.000000…1 kg. The cause of this change could

[142] Whilst he does not explicitly refer to the vase as a complex system, he assumes that complex systems (an organism) are capable of whatever autonomy a ceramic vase is. The notion that a vase is a system is relevant here because we have identified 'system' as necessarily individuated (often with an assumption of *integration*, as well). Most definitions of *complex* systems, however, include some element of computation, on top of structure, function (programme), and/or feedback loops. (e.g. Holland, 2006) The role of feedback loops relative to V above is intriguing because of a kind of environmental dependence (see §3.1.2).

be an observable change in the conditions around V,[143] but it could also be because of changes in the chemistry of V itself.[144] Conclusively, the mass' apparent lack of change does not prevent the mass from changing, and there is a certainty that the mass of an object will change—even if the fluctuation in mass originates from occasionally trapped force particles (such as photons or kinetic energy, which do 'carry' mass).

One of the foundations for endurantism is the positing of intrinsic properties, the most common example of which is mass. As discussed in Chapter I, intrinsic physical properties are no longer considered a given (even with spin and charge), as the issue is very much a contested one. When an object's primary intrinsic property is called into question, arguments for wholly-persisting objects should, in principle, also be challenged. These first points above on 'dynamic mass' suggest that no object maintains its mass through multiple instantiated observations. One way out of this challenge, of course, is to fix the number of decimal places in the calculation of mass at multiple instantiations.[145] The most common technique, though, for asserting stable and potentially intrinsic properties, is to exclude inconvenient levels of analysis, as we will continue to see in standard cases for MCP.

Additionally on composition, V is presumably composed of some kind of plaster, which has been moulded and coated with a polish. The plaster is held together over time by adhesive chemical bonds between the base powdery substance and (presumably) water. After the excess water evaporates, and enough bonds have set, the plaster hardens, allowing it to remain in a certain shape for an extended amount of time over suitable conditions (maintaining roughly the same shape that the moulder intended). The chemical bonds themselves also have much stronger bonds within them at the

[143] e.g. a loose particle of plaster or paint swept away by the wind, which had had a weakened chemical bond with the remaining constituency of V
[144] e.g. a delayed chemical reaction between the plaster and the polish applied to the plaster, which caused instability in V's structure
[145] We know that weight is relational or extrinsic, and that mass is generally considered to be intrinsic. But if the mass of an object changes between two instantiations, then how can one object be considered to spatiotemporally endure? How can any object have any instantiated properties that are intrinsic? This is to say: how can there be any intrinsic properties? This line of questions eventually leads to: what is an object, if all properties are instantiated, extrinsic, and accidental?

level of the atom and the nucleus.[146] So, the plaster in V is held together by a sufficient number[147] of molecular bonds, which are held together by atomic and nuclear bonds. Since the environmental conditions remain relatively[148] stable and suitable—as evidenced by the (continued) existence of a plaster structure—the structure seems to maintain a relatively constant chemical composition (which creates and maintains its shape). The hardened plaster's shape is to some small degree reinforced by the polish,[149] which also serves to: a) protect the predominant chemical component (plaster) from harmful substances (such as more water or corrosive chemicals); and b) reflect more light off of V than the plaster would by itself for aesthetic purposes.[150]

3.3.1.2. Environmental omissions

Gravity is the first obvious external omission, and in multiple understandings. V, which had previously been perceptibly[151] stationary at t_1, begins its fall toward the centre of the earth.[152] The pull on V by earth's gravity is a force which acts upon that which is contained within V, and not upon V as a whole. Whilst in the presence of vases, gravity does not behave according to 'vase laws.' A physicalist ontology requires that we explain events in terms of matter and force particles and fields, and thus we might say that V's densely-packed molecules (to choose one level of abstraction) cast the impression of a 'unified' object. This seemingly-unified, but presumably just tightly-

[146] e.g. in the strong and nuclear forces
[147] 'Sufficient number' is important because these bonds do not remain forever, and thus the object's constituents do not remain fixed forever. We can therefore say that V's constituents are not fixed.
[148] 'Relatively' is significant because the conditions appear to be sustained, but in actuality they are imperfectly consistent as a result of predominantly-undetected micro-level changes.
[149] Although, the polish's presence actually weakens the resolve of the 'structure' because—though intended in part to decrease the likelihood of slight 'chipping' in normal household handling—it actually decreases its frictional value μ. The impact of the polished ~1kg vase on a rough gravel or concrete, which has a very high amount of friction, would create a wider scatter of vase shards, along with polished and unpolished plaster particles.
[150] Considering these and other factors, our minds would reinforce the idea that the object is at least artistically whole and fixed. Even if V were to be only slightly cracked before t_1, just like a tiny scrape on a painting, we would consider these items to be the externalised and often tangible product of another's mind. This exacerbates the idea that the whole, as a piece of art, has autonomous causal powers, because our mind treats V as a unified whole casting an impression upon an observing mind predisposed to seeing (enduring) objects that exhibit minimal or insufficient change.
[151] The position of V is _perceptibly_ stationary because we do not observe how V is actually moving along with everything else on the planet (which is also moving in the warped so-called 'fabric' of space-time).
[152] The centre mass is the centre of gravity, so when objects fall to earth, they rest on the surface only because a denser object obstructs their path.

aggregated object is a concentrated hunk of matter and force particles (that only partially persists through to the next instantiated observation). The gravitational pull of the earth—which is a comparatively denser matter and force aggregate—is greater than that of V (due to the inequality of mass). Thus the object is pulled to earth not really by a singular, external gravitational force, but through *the event of* the battle of gravitational forces.[153] However one wishes to describe it, gravity is the primary reason V—or to be more, yet not perfectly precise, V's molecules—fall to the ground.[154] [155]

If one agrees that the nature of the impact of V on the ground is affected by the nature of V itself, then a thorough description and explanation of the nature of V is required before agreeing on the nature of the impact. The precision in any specification of the microphysical conditions of and around V cannot be overvalued; as such, tiny variations in precipitation and (subsequent) chemical reactions, for example, partly determine the nature of the impact. Kim wants to use molecular activity as evidence in one instance, but not in another, as if molecular activity ceases to play a role in the behaviour of V.

In addition to gravity, there are an infinite number of possible changes in the conditions 'around' V that: a) affect V's own parts or microphysical (chemical, molecular, and even atomic) conditions; and, therefore, by extension, b) allow V to fall to the ground. Why should one assume that the falling vase is not drawn 'by gravity' through a hail storm, a piece of which might impact V, and alter its mass, trajectory, velocity, momentum, etc.? What about a rainstorm, which would change the amount of friction both of V's surface and the surface which it encounters? How about an extreme condition such as toxic or unstable materials that radically change the chemical

[153] It is a very brief battle that ends in V's defeat. The speed of this battle can be calculated to a finite number of decimal places—depending on the object's mass calculated to a finite number of decimal places—with the gravitational acceleration for most altitudes under normal (zero wind/precipitation) conditions on earth being 9.8 m/s^2. Of course, the amount of time yielded from such calculations is also approximated roughly from central atomic clocks (also subject to standard earth-bound temperature and pressure ranges).

[154] If V was thrown, the momentum (Mass*Velocity) would be altered, as well as the trajectory, but gravity would eventually 'overpower' any kinetic energy transferred to V by its projector and draw it towards earth's centre mass.

[155] Gravity is important in the discussions of both the compositional and the environmental omissions. The incomplete understanding of the gravitational forces within the realm of sub-atomic physics is one of most glaring weakness of all physical theory, but establishing the nature of 'gravitons' still seems to be a concern for contemporary particle physics.

makeup and even mass, which would therefore alter the other properties as well? In some sense, any precipitation—such as a water droplet (more common) or a piece of hail (less common)—is 'getting in the way of V'. However, the converse could also be said: V is getting in the way of the rain droplet, and so they will affect each other upon collision (just as V and the ground are, as V is pulled towards the earth's centre of gravity). Ultimately, there is no limit to the possible events that V could experience during its descent, and our probabilistic calculations are of mere parsimonious utility. The inability of the human observer to imagine even a wide range of possible events involving V is largely due to our conception of and calculations within the relatively short duration of V's fall.[156] The spectrum of possible events involving a vase falling from a ten-story window, as opposed to V's two-story drop, would most likely be larger in our minds because of V's prolonged 'exposure to the elements'. But, at the molecular and atomic levels, both falls present an inconceivably vast amount of time for the particles to move, interact and aggregate.[157]

Regardless of these possible events and conditions, and whether or not we observe or predict them, I thus argue that we can say that the nature of V's impact on the ground is heavily—perhaps overwhelmingly—dependent upon the immediate environmental conditions. The author in this example has not accounted for their role in any of the ascriptions of powers and properties, and therefore it seems as though the claim for V's autonomy is incomplete.

3.3.1.3. Why V is not an MCP whole

As it has been insinuated, the path that V takes through the air molecules and the violent crash that ensues would only have taken place in that particular manner—amongst other, though less important, partial causes—if the force of gravity was sufficient.[158] Consider the alleged causal powers of V in zero gravity: the release of a

[156] Such a predisposition is also partially caused by our assumption that the environmental conditions around V will remain consistent or insufficiently dynamic.
[157] It is ironic that both falls present an inconceivably vast amount of time for particles to move and aggregate, because time itself for physicists is essentially the perception of atomic decay.
[158] And, in this case, it is approximately constant anywhere on earth, with the approximate same object released from the approximate same distance.

similar vase from a similar distance would have none of the same visible or actual effects. We would not use this example to demonstrate the "causal powers" of a vase, because they would not appear as obvious as the violent crashing of plaster onto concrete. Yet the explanatory gap is identical in both instances when erroneously supposing that a) V has autonomous causal powers, and b) V is an enduring object.[159]

With regards to a), under no circumstance can we claim that V has causal powers autonomous from its physical constituency. It has been shown that the path V takes is dependent on its parts, as well as its parts' interaction with its environment. Furthermore, and most crucially, V's existence depends primarily upon the conditions that immediately surround it. For b), temporal parts can only be excluded (from claims of endurance) if certain parts (and their temporality) are deemed extraneous. To claim that V is an enduring object and that only the extraneous parts[160] change or disappear, one would need to decide what levels of analysis to recognise. If one is interested in examining V in terms of its plaster grains, but not its molecular or atomic units, the argument would seem strangely anthropocentric and hugely inadequate.[161] Unfortunately for the endurantists, their philosophical position is immediately confronted with overwhelming contrary evidence that suggests the very composition of objects purported to endure depends on parts that objectively fail to endure in the same manner.[162]

Kim's V stands as the unit of analysis as he assumes V is a discrete whole—though, as I have mentioned in Chapter I, he does not provide any mereological criteria, such as whether the vase is something from which nothing or no part is missing.[163] Nevertheless, putting aside the dynamic constitution problem mentioned in 3.1.1, let

[159] To grant autonomous causal powers to V, Kim would need to claim V endures from at least t_1 to t_2.
[160] Extraneous parts would be considered inessential (Aristotle) or independent (Leibniz) parts. Endurantists today might consider trapped force particles as such, or having no ontological dependence on the whole.
[161] Of course, one can claim V endures simply by rejecting a physicalist ontology, but then the number of problematic claims would grow exponentially.
[162] The case for enduring objects can probably only be advanced in the context of intensional wholes, as the case of an enduring physical or extensional object seems problematic.
[163] Stating such criteria for whole qua object could have given his argument more clarity or heft. Additionally, I have already mentioned that distinguishing between these two definitions of whole is useful before a physicalist assumption has been declared. Kim, of course, is assuming standard physicalism (presumably with some kind of modeled primitive).

us look at the secondary problem of whether V is an MCP whole. The goal had been to explain how V's molecules were slaves to V qua whole (though not in Sperry's reflexive way), and that it was their existence within V that caused them to be scattered and redirected upon impact. The argument therefore had been that V's 'parts' or constituents (again, parts in Kim's argument) were the passive ingredients of an individuated work of art. We have seen his selectivity about compositional and constitutive causality and agency. There is another kind of error though, which brings us back to the lessons of Chapter II.

He acknowledges that V is a spatiotemporal whole on all levels, top to bottom, but he is unclear as to why he calls it a whole. Of course, V is a whole *prima facie*—though perhaps the better term is whole *simpliciter*, or without qualification. We can say *simpliciter* because it is an object we might roughly hold and place somewhere, with a roughly discernible shape, perhaps with some vaguely artistic features, and for many of the other reasons mentioned in 3.1.1. But these are largely for literary convenience, are they not? Since the stakes are so high—declaring that a vase has autonomous causal powers—we need to be as precise as possible. Is he saying there is some feature of V that makes it a whole? Our historical approaches generally identified a quiddity via essence. Now, he would not accept a unifying form or essence *per se*, but he may accept some other kind of quiddity. Perhaps this ascription of quiddity comes from these practical reasons just mentioned. However, his justification for treating V as a whole is his baffling choice to describe V as a system.[164] The phrase he uses to identify V's quiddity by extension is 'system-level properties', and this is of course a phrase I have used frequently in this very work. As we now see, the frivolous ascription and marshalling of system-level properties is varied and can be dreadfully misleading. I shall return to this problem in §4.0.

[164] Considering V as a system is perhaps a disservice to his own programme, because it means that any object similar to V is also a system. In some sense, sand castles, brown paper bags, and chocolate cakes would also then be systems, perhaps because they are objects with many parts and we are predisposed to think of them as discrete.

3.3.1.4. V as an integral whole

Let us build on conceptualising V *simpliciter.* I have acknowledged that V's quiddity might tenuously rest in our perception of V's topographical and artistic properties, though that would still not give us reason to ascribe essence or quiddity to it. Variable physical features and conditions relating to what are essentially topographically-contingent properties, such as weight, height, mass, constituency, temperature, and density are all *sufficiently insufficient* to justify that first argument. The second, 'artistic whole-ness' argument is that when we look at or create V as a piece of art, so the argument goes, we do not necessarily install a universal quiddity within it—but we do ascribe unity to it. V is not the same artistic whole to the artist-observer and all the general-observers—it does not have the same effect on both categories of observers. The same goes for any piece of art: no one would claim that a particular motion picture film has the same intrinsic qualities for all people—though this is largely because there could not be any intrinsic properties of such an object. Nevertheless, all general-observers, especially in a museum, will consider V to be a single piece and a single unit of analysis. Neither category can claim unity, however, because physicalist analysis does not grant access to a complete catalogue of sufficient physical conditions and properties identifying V as a *complete* individuated unit of analysis. Even the artist-observer, who in quite crude terms had simply assembled some loose bits of plaster—has no right to claim that it is something from which nothing is missing. He or she could not have had access to such a catalogue.

So with V as a non-MCP, seemingly disintegrated whole, where does that leave us with V's quiddity? As I've tried to show at least indirectly since Chapter I, no knowledge of V, no assemblage of its properties—no matter how seemingly exhaustive—will ever be sufficient to identify the essence of V. From the medieval recognition that some wholes have things whose loss the whole can survive, comes the realisation that V has things that do not necessarily contribute to what V might be described as. From a refined physicalist and empiricist perspective, we know, going further, that that which we do not know about V may not necessarily be that which is not required to be known about V in order to actually know what V is. In plainer terms, the synthesis, which clearly (and, as I see it, unavoidably) weaves both the ontological and the epistemological, becomes:

- *(1) V has some bits that come and go, or are seemingly less important than other bits.*
- *(2) We have not understood some of V's bits very well at all, so we cannot really be sure about how important they are to V.*
- *(3) It is possible that we do not understand V very well at all, either.*

As soon as we admit that there are some things within V that change, disappear, or cease to exist, I propose that we introduce a reformulated integral whole: wholes of which some things of allegedly lesser importance are missing—whose loss the whole may conceivably survive. Traditionally, the *totum integrale* still had an essence. Dropping any essential properties or disassociated form of any kind, yet keeping some kind of quiddity, we may yet still have a whole. I will present a full conceptualisation of the emergent integral whole in §4.0. In sum, V can only be seen as an aprioristic MCP whole, because we could not qualify its integration nor justify its individuation using our refined physicalism and empiricism.

3.3.2. Case 2: The Street Market

Recalling his anti-reductionist contributions mentioned in Chapter I, let us revisit Dupré's piece on "The Constituents of Life", where he has tried, as he has often done, to convince the reader of the insurmountable gap between what are usually biological, ecological, and cultural agents, and the cold, lifeless, physical emptiness and non-structure that we find apparently everywhere else in the cosmos—including in particle accelerators. His example of the flowery radiance and bustle of a popular street market (M)—contrasted with the dull metal and plastic of an automobile—is apparently intended to show that M has the ability to create change in the world. The author is, of course, not declaring in this example any particular ontological, mereological, or epistemological foundation from which the claims can be inserted neatly into the metaphysics of emergence, but I do believe that the following example, which is indeed contributing to the philosophy of science, makes the same type of mistake as several of the authors mentioned in this work. [165]

> [M] has an unusual capacity to attract people, a capacity which, I suggest, has significant similarities to the ability of a flower to attract bees, or the ability of a magnet to attract iron filings: all are causal powers of individual things. The particular causal power of [M] will be obvious to anyone wandering around the streets in the immediate vicinity: while there will be a light scattering of people in these surrounding areas, immediately when one reaches [M] one will encounter a dense throng. The reason is no mystery, of course: this is a busy street market. The market could not exist without the people (and stalls, and products) that make it up, but equally there are properties of the market itself that attract the people to it.
>
> The powers of this market are exactly matched to the powers of the people it attracts. They must know it is a market, for instance, and how to get there. These are not difficult accomplishments: I myself managed to acquire them within a few days of arriving in [the city]. But of course I had acquired many of the necessary skills years ago: knowing what a market is, how to buy things, and so on. My return to the market to forage after my first accidental encounter with it is, however, a more complicated achievement than, say, returning to a place where I had previously discovered edible berries. I would not be similarly drawn to return to a place where I had seen delicious looking food through the window of a private house, for instance, and I would not return to the market at four o'clock on Sunday morning. The market is a social institution of a kind that I have learned to negotiate reliably. By learning this I have also become – willingly, I should add – susceptible to the attractive casual powers of this institution. The market depends for its existence on the people who go there to buy and sell; but it is simultaneously the power of the market that attracts the people that constitute its continued existence. And, insignificant though these may seem, the market effects changes in the people it attracts – it may determine, for example, what they eat for dinner. This is the sort of thing that I mean by a node in the causal nexus. I shall suggest that this model, incorporating the development of two-way causal interaction between a complex thing and its constituents, is the right model for interactions at many different levels of structural organisation. (2008, 15) [166]

[165] Many of course will wonder how this approach can be contested. But, when looking at the metaphysical implications of such a claim or approach, we see that it presents a massive problem: what we find in accelerators and stars and inside atoms is exactly the same things we find inside these same objects. The question has always been about the relationship between levels.

[166] It should be noted that this piece was initially a lecture (though later published as a book), and it may be unfair to scrutinise so closely that which was meant, in part, to be heard, and not read. However, I am considering the words as they have been written and published, and am assessing as I would any written claims.

3.3.2.1. Initial concerns

First, let us look at what is implied by ascribing autonomy—or M-ness—to M *simpliciter*. M is essentially identified as a topographical, three-dimensional object[167] with irreducible, autonomous causal powers. Presumably M is located somewhere with at least approximate physical boundaries. Perhaps one can say that M is 'contained within' a particular square between certain streets 'operating' between certain times on certain days. But, these descriptions of M are easily challenged as evidence of M's autonomous existence considering their imprecision and ultimate parsimony.[168] And, are these vague, logistical properties the only properties that we can ascribe to M?

The author is arguing that the human observer Q is *affected* by M qua whole, and not by that which is contained within M. After all, as the description states, Q is interacting with M, not M's parts or constituent objects and events and their properties. Yet curiously, the autonomous system-level properties of M are never actually identified beyond the vague topography already mentioned. We learn that nearby pedestrians observe heightened density and volume in a roughly-defined space, but what properties of M can actually be individuated, especially at t_1? A memorable experience of visiting what we have been calling M would need to include references to specific stimuli, for a declaration of enjoyment of M would normally require more detail: Q's observations and 'experience' from t_1 to t_2 is processed via properties. Q would be prompted to mention the delicious tea cakes or the hand-woven turtleneck sweaters, which may very well have caused Q to regard her experience 'in M' from t_1 to t_2 as enjoyable. One could also suppose that Q is allergic to, or is disgusted by, tulips and trout, and the closest two stalls offer exactly those things. The fish smells and pollen in the air outweigh Q's curiosity in the third stall, which happens to be offering truffles.

[167] If the author's M is not three-dimensional or an object, his argument becomes even weaker and has even more work to do to justify his claim of M's autonomous causal capacity.
[168] Which calendar are we using? What measure of time are we using, and whose clocks are we following? Is there a point at which I am considered to be "at the market" and "not at the market"? One unapproved vendor may be several metres away from the concentration of approved vendors. If I peruse his stand, am I at the market? The list of these seemingly trivial points is endless, but they are meant to illustrate the vagaries of the alleged properties of M—and therefore identify the tenuous nature of a claim for M's autonomy and whole-ness.

It is clear that there may very well be a large number of possible stimuli topographically *associated* with that which we have been calling M, but only in Q's mind might these stimuli be ontologically *contained* in such. The author wants the tea cakes and trout to be properties of M, but alleging that they are—as tenuous as that is—is insufficient for considering M as an autonomous, causally-efficacious whole.

3.3.2.2. Further extrapolation

Is there not a possibility that similar stimuli exist from t_1 to t_2—or had already existed—elsewhere? Are the individual stimuli not available anywhere else? If they are available elsewhere, it is by pure randomness that they appeared in M *simpliciter* from t_1 to t_2. Furthermore, Q's impulse to investigate what appeared from a distance to be tea cakes might very well have arisen from memories and experiences of tea cakes. More broadly, street markets often arouse our curiosity in general possibly due to their presentation of fresh or hand-crafted goods, for example. We know there is a chance that something of some interest may exist within the market. We do not approach a street market because we want to enjoy the market; we approach because we might find something interesting 'within' 'it'.

It seems as though we are meant to assume that all observers will flock to the market almost predictably, like iron filings to magnets. Yet, surely it is conceivable that the beauty and wonder of M is not irresistible! Surely he would not be implying that Q's attendance and/or participation was inevitable. Magnets—though nowhere near as complex as the human brain inasmuch as they are both individual units of analysis—are not really the exclusive macro-level source of agency in their attraction of iron filings. The attraction itself is contingent upon, amongst other things, the ionic bonds present in the chemistries of both the magnet and the filings. And, surely the author is not claiming that it is certain that magnets of any kind will attract any and all iron filings.

In placing these component stimuli under the umbrella of M *simpliciter*, are we not obscuring the highly unlikely convergence of so many other objects, events and properties? Impossible as it is to calculate a probability of such, is it not deeply impressive how unlikely the odds are that Q happened to walk near those stalls at that time under those conditions? Additionally, some may claim that when we take qualia and intensional wholes—e.g. ideas—into account, we may begin to not only ascribe unity to what are otherwise fragmented units of analysis but also to grant autonomous causal powers to them. A non-spatiotemporal version of M, proponents might say, has macro-level powers—the ability to instantiate new properties in the world, such as new mental events and/or physiological responses to them. But, that version of M would have no connection to the physical components of M if it is indeed non-spatiotemporal. For our purposes, though, the physicalist had never granted any concessions to such abstract objects.

3.3.2.3. Other similar non-examples of MCP wholes

If one was looking out over a rugged, high-altitude, mountainous landscape, perhaps with thoughts and feelings of inspiration or refreshment, do we say that the mountain has affected us? Does the mountain have autonomous causal capacity, or the ability in and of itself to instantiate new properties in the world? In a very poetic sense, it can be considered a whole, a singular unit of analysis, from which nothing our minds can consider is missing. It is the mountain M, whose quiddity includes system-level properties such as approximate height, diameter at sea level, and cartographical location. Again, the topographical argument is vague and insufficient, and the claim that the *idea of M* is an MCP whole, capable of forging new qualia, is irrelevant and ultimately untenable. So how could someone claim that the mountain effects changes in them? I believe at least the majority of observers prefer or subconsciously succumb to clustering stimuli together, as with so many larger-scale settings, events, macro-objects, and anything we consider ourselves to potentially be a part of.

We see much of the same with Q at a large sporting event. When Q claims to be ‘uplifted by’ or ‘enjoying’ the unit of analysis M, called the ‘spectacle of the event,’ it is an injustice to the full spectrum of component stimuli present ‘within’ M *simpliciter*. The loud music, the screaming fans, the action on the pitch—these things are reducible to and explicable by other things—and they may not be metaphysically associated with whatever Q thinks M is. The spectacle that we are calling M is not Q's environment *per se*, nor is it a discrete object or event in Q's environment, whether it be at t_1 or from t_1 to t_2. M is a vaguely identifiable, yet abstracted, highly-selective and grossly incomplete collection of properties within Q's environment. More precisely, M is a *substratum* (see note 118) of the environment of the particular, a particular Q in this case that is capable of computation and detection of what we know is quite a small range of potential stimuli within that substratum. The particular gathers the data from the stimuli and associates them collectively (and very often instantaneously) with that substratum. Q associates the truffles with the truffle stall, the trout with the trout stall, the tulips with the tulip stall, and the truffles, trout, and tulips with M—and that association has no metaphysical significance.

3.3.2.4. M as an integral whole

From the examples above, at no point should we de-value the individual stimuli; for only in us as individuals with cognitive faculties do we assemble their additive and interactive effects, properties, qualia and mental events they may happen to trigger. But, is there some way we can salvage M as a *quidditious* whole *of some kind*, so that we might have a unit of analysis with which to work? We obviously cannot claim that M is an integrated object, because it is physically *disintegrated*. But, in what sense and by what means might we be justified in individuating M qua object? How can we say that M has a quiddity?

Just as with V, the claim of a unified M—from which nothing is missing, and of which there is an autonomous causal capacity—is unjustified. M *simpliciter* then becomes something to which we can ascribe some properties to in much the same way that we slap yellow adhesive labels on common objects. What is written on the label is a scribbled word about or brief summary of the contents of the object to which we have applied it. Suppose that M has been placed into a box, as large as it needs to be, where the box itself is the shell of M (and therefore M, too).[169] M's box has quite a few labels, but it would be foolish to claim that all possible labels have been applied to it.

[169] This should remind us of the physicalist spatiotemporality required by our present analysis.

Some of the labels have been changed or removed between t_1 and t_2, and some of the labels refer to aspects of M that many other non-M objects also have. Furthermore, some of M's labels are really labels meant to be applied to something inside the box that nothing else inside shares, and one might feel at least slightly uncomfortable putting it on M's box and not on that something inside it. As we look at M, resting in a box from t_1 to t_2, covered in yellow labels, we may be having a hard time understanding M's quiddity and maybe even questioning if identifying one is possible. Is it? And, if so, which labels, or, perhaps, which *collection* of particular labels, mean something metaphysically significant? Should we even consider the possibility that all objects have quiddities?

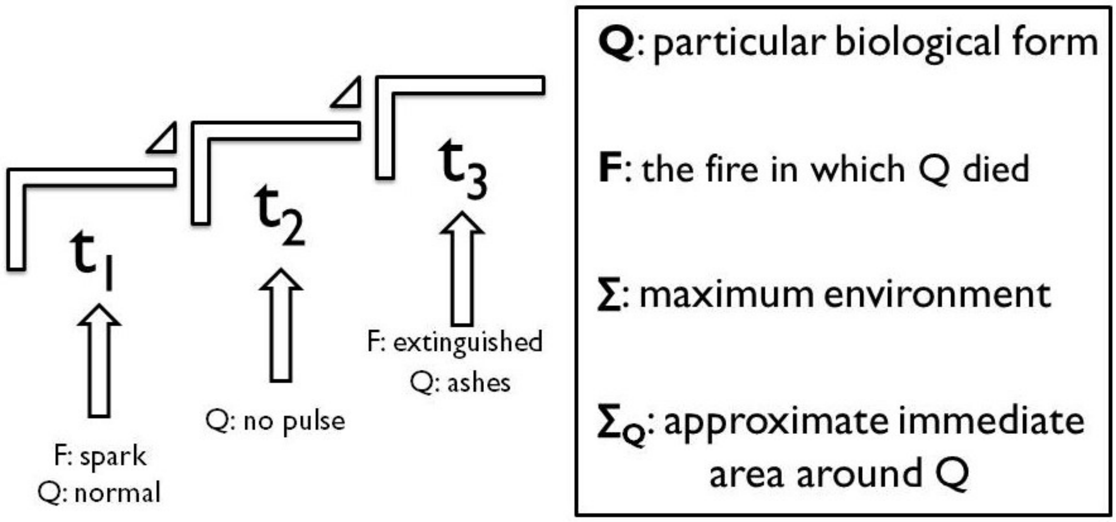
I am gradually approaching what I hope is a much more helpful synthesis of the analysis I have presented thus far. Just as with V, though the claim for M's reflexive autonomous causal capacity is difficult to make, there is still some hope that M may still be considered a whole—just like the vase, the automobile, brain, the cell, the molecule, and the tree. As we will see, physical integration is *not* a necessary criterion for whole-ness qua unit of analysis in a physicalist programme. Some things—of varying degrees of importance to whole ascription—do not need to be present and continuous in the whole *simpliciter* for it to be considered as such. And, most importantly, whilst there is no room for a metaphysically-significant MCP, the quidditious integral whole *simpliciter* might be seen as *emerging* as a dynamic and diachronic substratum of a particular's environment.

3.3.3. Case 3: The Fire

The following example features a whole that is perhaps more difficult to conceptualise, yet made of the same kinds of basic physicalist constituents as V and M above. It is also a more powerful example of the role of the 'environment' in any ontological and etiological discussion, particularly when MCP claims are at stake. Whilst it will prove difficult to see the fire as an object, and that it will seem to be more of a case about causal powers than wholes or objects, the purpose of this example is demonstrate the difficulties in ascribing autonomous causal powers to objects that we normally struggle to individuate. The struggle, it seems to me, should be equal to what we experience with typical wholes or concrete objects.

- *A biological form Q dies in a house-fire, and all its personal affairs—including several pets and houseplants—are destroyed. The initial spark[170] occurs at precisely t_1, Q is dead and lacking a pulse (under standard measurements) at precisely t_2, and the last smouldering ember has been extinguished (with Q as pure ashes) at t_3.*

Figure 3.3.3.1: Tracking the Fire



Can we say that the fire in which Q died—henceforth F—killed Q? In the English language we do not normally use the noun 'fire' with a transitive verb, nor even do we typically use it in the nominative case in the active voice. It is not grammatically incorrect to say that "the fire killed him," but it is more likely that we would say that "he died *in* the fire"—which translates into "Q died in F." But, that would imply that an event (Q's death) takes place *in* a space (F), such that F could not be an object or an agent. There is, at least in conversation, an obscured sense of F as something that is both a region of space-time as well as an agent. It is conceivable that this problem is a result of imprecise language, but perhaps not—as we may see.

Nevertheless, does it not sound awkward to say that "a fire killed Q?" We might say that a series of forest fires has killed off many species of trees, but the sense and grammar has changed. If we wanted to know the Humean cause of Q's death, we

[170] This is the initial spark of the fire which we come to recognise as 'the fire in which Q died', or F.

would probably say “the fire killed him,” which translates into “F killed Q.” Thus, the same question is asked of the fire that was asked of V and M above: does F have an MCP-like agency, or autonomous causal capacity? [171]

3.3.3.1. Examining the question

For F to be an MCP whole, it must first be a spatiotemporal object. In our minds, perhaps, it would seem to be much easier to claim that V, a common object, is such a thing, even if only at t_1 or t_2. M is similar to F in this regard, as it not something that we can, as it were, place on a shelf—or under normal circumstance, put into a box. It is clear that in order to be a whole, MCP or otherwise, it must be an object. So, why is it so difficult to consider F as an object?

We know that F is “the fire in which Q died”, which takes place from t_1 to t_3. The yellow adhesive label that we had put on M’s box is not so easily ascribed to F, especially considering the dramatic changes of whatever F is from t_1 to t_3. Looking back, we can see how it would seem to be much easier to claim that V’s labelling is adequate, even if only at t_1. Perhaps V can be said to be an object *simpliciter* in virtue of its comparatively and perceptibly stable chemical conditions. Describing F though, even at t_1, is very problematic, largely because of its *plasmic*, volatile, seemingly-uncontainable state. F is composed of things that change—so, F changes. How can we be sure that F at t_1 is still F at t_3? We will need to be sure of this in order to claim that F killed Q, because Q did not die at t_1. Q was pronounced dead, as Q lay as ashes, at t_3.

Any speculation that F may be an MCP whole needs to address the following questions:

[171] There is an emphasis on the following statement: “the fire which killed Q” is not the same as “the fire in which Q died.” Granting causal powers to the fire as a whole is far more problematic than simply identifying some of the properties of the space in which the object is (momentarily) located.

- *(1) Is F an object? → Does F have physical or spatiotemporal properties at least at t_1?* [172]
- *(2) If yes, does F persist? → Does F have at least some of the same spatiotemporal properties at t_1, t_2, and leading up to and/or including t_3?*
- *(3) If yes, did F kill Q? → Is F an MCP whole?*

(1) F, like V and M, has some loose collection of spatiotemporal properties at t_1. It has approximate boundaries: F is contained within a house or room, and not in the house next-door or on the street outside. Smoke and flames are detectable, as well as higher-than-normal temperatures. The existence of a spark is dependent on the existence of oxygen and suitable conditions for combustion. One may identify a roughly-contained region of the house where F started. Ultimately, there are some generalised, crude properties of F at least at t_1, so F is loosely considered an object at t_1. (2) Some of these properties at t_1 are discernible at t_2 and possibly leading up to t_3: it is still contained within roughly the same region of space, there is evidently still some presence of oxygen and suitable fuel for combustion, and the smoke and flames—though wildly intensified in size, heat output, and colour—are still present as general system-level properties of F (again, we will return to system-level properties in the next section). Again, ultimately, there are some generalised, crude properties of F that can be identified at both t_1 and t_2, and hence, we can say that F persists in some way.[173]

(3) In light of the hesitant admission that F is an object, that persists in some way, one might be tempted to consider F as being an MCP whole. But, we never did feel comfortable saying that F killed Q, so we were never fully confident that F is an agent—having autonomous causal capacity. Our initial definition of F was "the fire in which Q died," but that does not require F to be an agent. It merely implies that F could also be an event, or a region of space-time, in which Q co-exists, participates, and ultimately perishes. So despite the affirmative yet tenuous statements about (1) and (2), there is still apprehension about (3).

---

[172] I mean here to ask whether any idea of F includes properties exclusively as a result of it existing in time and space. The notion that F is 'interesting' or 'humorous' has no correspondence with F qua *res extensa*.

[173] The reader is reminded here that it is acceptable, in my view, to assimilate F qua whole with F with MCP because both would need to be integrated and individuated.

Going back to persistence, there is no necessary compatibility between the claim of an object loosely persisting and that an object with outwardly-facing autonomous causal capacity. All we had to satisfy (1) and (2) were a rough collection of properties ascribed to F at multiple instantiations. There were enough of the same yellow labels applied to F at t_1 and t_2 to call F a persisting object. In order to subscribe to (3)—to say that F killed Q—one would need to have more than an admittedly biased and possibly incomplete range of properties.

In fact, for F to be an *autonomous* agent, F needs to be an *enduring* object. This is the unfortunate and troubling realisation of any structure-agency debate: in order to claim autonomy in agency, one needs to claim either essential properties (a claim which could easily be disputed) or a complete catalogue of properties (a claim which is easily dismissed). The former involves essential wholes, where the observer—let us say, in this instance, Q is the observer, for he was already there—has ascribed an immutable, possibly even non-spatiotemporal quiddity. The latter involves enduring wholes, where the observer has claimed a whole as something from which nothing or no part is missing[174] and has persisted without loss or change.

I have largely favoured essential wholes in favour of integral wholes, because of the inevitable compositional volatility and non-spatiotemporality of *essence*. In addressing the issue of enduring wholes, I would ask if all properties of F remain consistent from t_1 to t_3.[175] A parsimonious approach would claim that we know *enough* about V (from §3.1) at (at least) t_1 to merit consideration of V as a) a spatiotemporal object, and b) roughly the same object observed at t_2. However, the proposition that V endures—which requires that V's constituency remains intact from t_1 up until t_3—is untenable in a refined physicalism and empiricism. For F, this is an even bigger challenge. Again, for F to endure, we would need a full set of properties at multiple instantiations. In the case of F, we only have a very limited set of extrinsic properties at any of the

[174] Again, the reader is reminded that we should by now be no strangers of the weaving of the ontological and epistemological—and should even expect it particularly when properties are readdressed.

[175] The overwhelming, obstinate problem for endurance, though far from being the only problem, is that we know, at least at the quantum level, that particles come in and out of existence without our naked perception. We therefore know that in some seemingly insignificant sense F does not remain perfectly consistent. This is the same problem as with the trapped force particles in the vase, and the component stimuli on whatever scale 'within' M.

instantiations, and these sets are not valid for determining F's endurance. In the case of V and M, we have the same inadequacy of any individual set of properties at any instantiation, albeit to a lesser degree. Again, it might seem that V remains V at t_1, t_2, and leading up to t_3, but the sets were never complete.

For biological forms in a physicalist ontology, such as Q, endurance (total persistence) would require its material composition to remain fixed at minimum from t_1 leading up to t_2. This is immediately problematic, as one could easily point to various microphysical changes—even excluding those directly influenced by F—such as the exhaled carbon dioxide. In order for F to endure, it would need to maintain the same temperature, colour, size, smell, and a range of other properties. Clearly, the fire's average temperature—which, again, is a mean measurement of a mean measurement of an approximation—increases as combustion increases.[176] The spark is not the same as the roaring blaze which consumes the house. Though F's colours may oscillate minimally between shades of yellow and orange (though blue and white parts of common, open flames are usually present as well), what the human eye perceives will vary from individual to individual and as only an imprecise fraction of the electromagnetic spectrum. F probably would not survive beyond the spark at t_1 if its size remained the same—and this is only using a general, voluminous, spatial notion of 'size'. With regards to smells, we know that certain substances objectively emit particular odours, but only under similar conditions. If the house-fire's fuel is predominantly wood from the walls or floorboards, for example, it will give off a burning wood smell.[177] But, the smell of F is dependent on the observer, especially considering that the location of the observer Q near F matters. He or she may be standing next to a pile of decomposing polyester fabric, the odour of which being drastically different from that of burning dried wood. Ultimately, these properties that we had ascribed to F at t_1 will not and cannot find themselves necessarily instantiated at t_2 or t_3. F does not endure, and as we now discover, F is not an MCP whole—or, F did not kill Q.

[176] Temperature is a stunning example of an interactive property, as it is the measurement of the mean kinetic energy in an approximate region of space-time. It is a mean (T) of a mean (kinetic E) of a mean (region/field).

[177] The reaction is similar to burning sugar, where the reaction products are carbon dioxide and water. The 'smell' of such a reaction, though obviously classified as qualia, will be similar under standard room conditions.

3.3.3.2. F as an integral whole

Whilst its properties undergo change, F *simpliciter* is alleged to persist from t_1 up to t_3. This means that whilst we cannot grant MCP status to F on the grounds that it does not endure, there may be something left of F qua whole to salvage. If we consider F as an integral whole vis-à-vis F *simpliciter*, where F is a loosely assembled approximation of a whole—we can claim first that there are some parts and properties of F that stay and some that go between t_1 and t_3. There is some set of properties that allows us to see F in a narrative fashion—to see similarities between observations. Something about F at t_1 was there at t_2, and something at t_2 was there right up until if not including t_3. This is related to the worm view of perdurance that I will address in Chapter IV. In terms of properties, we see a continuation of vague sets of parts and properties linger between observations such that we can say with some reticence that F at t_1 has something in common with F at t_3. Because of this continuation in the weakest possible sense of continuation, we, as we have since the beginning of language, may ascribe whole-ness to F. But, there is so much more to say about F, as well as any whole with some elusive set of narratively-persisting parts and/or properties.

The *totum integrale* recognised both essential and inessential parts, where the loss of the independent parts the whole could survive. If we were to convert the essential properties into persisting ranges of extrinsic[178] and relational properties, and inessential and/or non-persisting parts as necessarily part of any object, we may be converging on a reformulation of not only the integral whole, but wholes in general. We know that there may be a range of smells, a range of colours, a range of sounds, and, amongst others, a range of temperatures that, at least for the observer Q, F stays within. It is not without some concerted effort that we can claim with some scientific confidence that F stays within a range of properties. We know obviously that, for example, F cannot and does not exist below or above certain temperatures: F, like the

[178] 'Extrinsic' here is once again meant to imply properties derived from the object's relation to another object. As will be shown, F cannot be said to maintain intrinsic properties, such as mass, from t_1 to t_3. However, I might argue elsewhere that there might be no need at all for intrinsic properties, if one considers that no spatiotemporal object has enduring, unconditional properties. If there are no enduring unconditional properties, or no ontological independence of some sort, how can there be instantiated or enduring intrinsic properties?

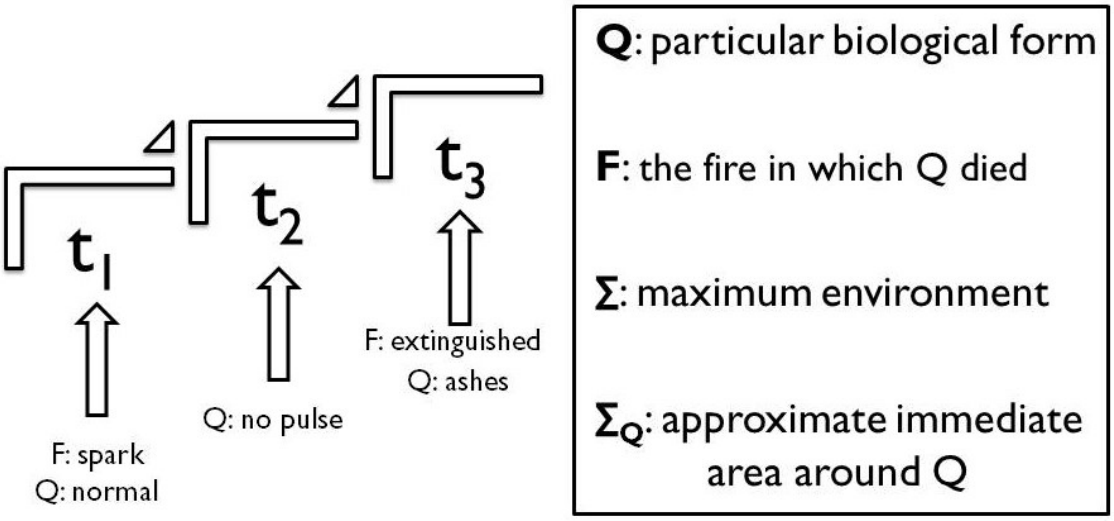
house in which 'it exists', could not have appeared if the conditions approached absolute zero on the one extreme or the core temperature of an average star on the other. There are some conditions that we should be prepared to tentatively exclude as possible conditions of F, on the grounds that their existence would prevent F's. But, since the range of conditions for the parts and properties of F that we observe and ascribe meaning to is still infinitely large, and since we never had a complete catalogue of what F actually is, we are required to:

- *(1) abandon the consideration of F as an essential and enduring whole*
- *(2) realise that any set of properties of F at t_1 is at least weakly dependent upon the observer at t_1* [179]
- *(3) realise that any set of properties of F at t_1 is also contingent upon the conditions around and within F at t_1*

The implications of (1) have been addressed. In (2), we are reminded of the limitations of human and human-programmed detection and observation. I believe (3) is the most intriguing, because it might provoke an alternative view on whole construction on a refined physicalist and empiricist foundation.

[179] This notion is something I have been discussing from the beginning of this work. At its core, it is *the* key feature of what I have been calling a refined non-linear, non-ontologically reductive, irreproducible, open-system physicalism and empiricism. I have provided definitions and included many examples for why there cannot be a complete catalogue of observation-independent properties. It is central to the thesis, and is a fundamental lesson from quantum mechanics and the so-called *anthropic principle*. Even more discussion follows in §3.3.4.

Figure 3.3.3.2.1: Tracking the Fire (repeat)



3.3.3.3. Distinguishing and identifying F

Being able as we are to identify only a fraction of F's properties, and place them in ranges of properties, which are then considered to be present at each observation, we have clumsily come to the tenuous conclusion that F may be considered as an integral whole, or F *simpliciter*. This understanding of F is so far removed from the *totum essentiale* that we might be inclined to say that all object identity has, as a result, been made nominalistic, parsimonious, instrumental, and asymptotic. And, I hope this is how it is to be understood. The reasons for this shift are most vividly apparent to me when addressing the nature of the relationship between wholes and their environment.

I began §3.3 by asking whether F had autonomous causal powers—whether F is an MCP whole. After safely showing that this is not possible (because of F's absent integration), I then moved to question the notion that F is even a quidditious whole. F *simpliciter*, as we have come to refer to it, is exclusively a loose collection of properties arbitrarily ascribed between observations. When we stretch F *simpliciter* out to multiple instantiations, in a weak attempt at proving its persistence, we have a difficult time re-assembling and re-ascribing any of the properties from earlier observations. Part of this struggle is due to the fluctuations in the conditions around the (let us say, mammalian) observer Q, to the point where F is difficult to distinguish from the area around Q. The fire as an object, let alone as a whole of any kind, becomes a tenuous proposition.

What, then, is the difference between F, and the environment (∑) of Q (henceforth $\sum_Q$)? It seems obvious that F is not the same as $\sum_Q$, because there are properties of $\sum_Q$ that could not possibly be properties of F—such as Q, Q's clothing, the walls of the house, etc. If F is not the same as $\sum_Q$, then what properties does F have? We recognise that ∑ is the environment of a particular, hence $\sum_Q$ is the environment of a particular Q. But, ∑ without a particular is everything that exists. There is no set of properties that describes ∑ at any instantiation. In other words, we can never fully describe, or have a full set of properties for, the cosmos at any moment, let alone through time.[180]

It seems to me that we must consider F purely as a substratum[181] of $\sum_Q$: it is an inconsistent set of conditions and properties from t_1 to t_3 in an approximated region of space-time. The notion that F is an object is in doubt because it lacks specific boundaries, autonomous or macro-level features, as well as a programme or intrinsic function. If F is just a substratum of $\sum_Q$, it then follows that F is an accident of $\sum_Q$. It simply happened that F was approximately in the same region of space-time as Q from t_1 to t_3. When phrased like this, it seems there would be many objects, events, and properties that would similarly be considered as accidents of $\sum_Q$. This is most certainly the same problem we encountered with M above: the trout and the truffles are not necessarily or exclusively dependent on M for their existence, so we may not claim their properties are part of (and therefore properties shared with) M.

How specifically can Q be affected by $\sum_Q$, which is a part of ∑? In the case of F, Q (as well as the houseplants) suffered a series of critical burns following from the combustion of oxidised materials and the subsequent release of intense heat. The sustained heat 'from F' caused overwhelming damage to the cells in the body of Q

[180] The irreversibility of the universe, and thus the indeterminacy of it, is an issue with an insufficient amount of discrepancy, and thus we proceed. The reader is encouraged to see Prigogine, Davies and Krauss for more on this.
[181] Again, I am using *substratum* in its original Latin sense of "that which is spread under". I am using it as one given perceptual plane: the vaguely-individuated region of space-time that an observer consciously or unconsciously recognises or believes is perceptible in some way.

that initially restricted movement and mobility. Intensifying combustion with oxygen as the primary fuel began to reduce the air supply for Q and other biological forms leading to a scarcity of oxygen. Q becomes unable to breathe properly, preventing its internal structures from receiving necessary resources to continue normal function and mobility—in a sense, preventing his escape. Unable to move or breathe, Q collapses and falls victim to intensified oxygen depletion and electromagnetic activity causing direct harm to the physical body. Eventually this all led to the decomposition of all biological forms and their physical constituents.

Thus it can be said that in this example F and Q cannot co-exist for very long if occupying the same approximate region of space-time. One can indeed place one's hand onto a flame for a brief moment and still survive. But, when the exposure is prolonged, the biological forms will break down. Under standard Earth-bound conditions, biological forms to varying degrees flourish. If local conditions deviate enough from a suitable range, such as a prolonged exposure to intense heat or density, the forms will break down. Thus, biological forms such as Q and the houseplants can exist only under certain conditions or a range of conditions. For example, we can say that certain biological forms can exist in Arctic or Saharan conditions, for those forms have developed and maintained suitability to those conditions. But the gravitational forces, temperature range, and density, for example, are relatively constant. There is some variation in the amount of kinetic energy trapped in these regions, but the variation is minimal on a cosmic scale. Now, obviously none of these latter points are terribly shocking, nor are they recent revelations. However, these points do lead to the conclusion that biological forms are ontologically dependent and causally contingent on a spectrum of other, entirely physical and local factors.

The observer Q is predisposed to seeing the flames across and between observations as a singular unit of analysis, if not as a whole. For the observer, this whole *simpliciter* can be extinguished, reduced, re-directed, contained, manipulated, and even excited. Ultimately though, for the average observer, it is an object, just like any other, that interacts with the world as an autonomous agent. Especially in the 'heat' of the battle between Q and F, F is an opponent to be conquered. The battle is not between Q and an enflamed piece of house furniture among many, nor is it a battle between Q and a few excited electrons. In the mind of Q, it is Q versus F. But, this cannot be the case, as we have just shown. The volatile molecular and atomic activity, and its wild, seemingly unpredictable behaviour, present in what we have been calling F, is too important to F's quiddity to call it a singular agent and a whole. The vague collection of properties, or more precisely, the collection of the ranges of various properties from t_1 to t_3 makes F exclusively an arbitrary and metaphysically non-unified part of Q's environment. $\sum_Q$ includes F *simpliciter* at t_1, where F *simpliciter* is one substratum of an infinite number of potential substrata of $\sum_Q$.

There are some simple conclusions to be drawn on the difference between F and $\sum_Q$. First, we cannot consider a substratum of our environment as an object or whole. The properties of a given substratum can never be holistically attributed, or ascribed, to that substratum, because the substratum was only arbitrarily defined in the first place.[182] We must acknowledge that some if not all of the properties of the substratum, such as $\sum_Q$, can be reduced to the properties of its parts.[183] This is the overwhelming factor with both F as well as M. The so-called *noumenal*, qualia-based properties of a given substratum, like the "roar of the crowd," "the feel of the market," or "the heat of the fire", are all trivial, subjective, intensional, and potentially dualistic property-effects of the mind assembling the micro-level physical properties and associating them with one whimsical selection of infinite possible substrata. Irrespective of the arbitrary mental assembling, F as an integral whole (because some parts come and go) *simpliciter* (instrumentally individuated) does not endure (failed integration between instantiations) and does not possess autonomous causal capacity (because its agency is disintegrated).

[182] This is equally as arbitrarily-defined as the broader $\sum$. The distinction between $\sum$ and $\sum_Q$ can only be meaningful or useful for Q.

[183] Property-ontological reduction is the real aim of all physics and most natural science, and this must be contrasted with pure ontological reduction. It is a fantastic misapprehension that all forms of reduction have failed, and that all forms of reduction inevitably imply downward, vertical explanation of entities. Property-ontological reduction is "step 1" of all science. Pure ontological reduction failed because it considers static macro-level entities as vertically composed of static micro-level entities.

3.3.3.4. The necessary extrinsicality of properties

Ascribing system-level properties to any of the above objects (V, M, F) has proven to be problematic in two senses, the first of which relates to the integration and individuation of the object qua whole. This is to echo one of the central questions of the thesis: how can I ascribe emergent property R to that which does not seem to be unified? This does not bode well for intrinsic properties, as demonstrated in this §3 with mass, spin and charge. The other sense has been floating in the background since the beginning of this work: does the extrinsicality of all properties not imply inherent context- and observation-dependence in property ascription? In the introduction, I defined a property as ‘coded sensory data’; so, ‘redness’ is the property given to anything that satisfies our shared criteria of red. Consider the following experiment: The members of Group 1, who sit in a lit room, all agree that an apple A is red. The members of Group 2, who sit in a dark room, cannot be certain that A is red. Group 1 ascribes the property red to A, whilst Group 2 does not. Does it make sense to say that A has the property red? The important point here is that if we were to say that A is red under normal conditions, we expose our *assumption* about normal conditions. The fact that A has a different appearance depending on local conditions means that the instantiation and observation (detection) of the property red is fundamentally dependent on local conditions, too.

I have already discussed at length my arguments against the most traditional of observation-independent properties: mass. I will restate them here to draw this section to a close. Mass is ultimately just a measurement of some magnitude of matter, and as such, it really only makes sense conceptually in relation to force. As with all magnitudes, units were invented (and agreed upon) to measure ranges of magnitudes, and all measurements are restricted to a finite number of decimal places. So, in order to have an observation-independent measurement of mass, we would need to agree on how many decimal places we wish to calculate to—which would necessarily be an approximation (multiple investigators would need to agree that this number of decimal places was adequate). And, of course, we would also need to be sure we are measuring the exact same object, which is precisely the central question of this thesis. We have a measurement of mass, but to which object shall we apply this property?

3.4. What is a whole *simpliciter*?

I hope my analysis above has not for the reader insinuated that there is no need to salvage some kind of quiddity, despite the disunity and loose persistence of the units of analysis. As mentioned in the opening section, we need to study *objects*, *events, properties* and *conditions*, and there cannot be investigation without units of analysis. If the wholes that we have been looking at—as well as the wholes that we will look at in the coming sections of this work—lack physical integration, or cannot be considered autonomous, MCD and MCP agents, then what are we to study? As I have tried to slowly unfold, recognising integral wholes (wholes we recognise as having parts whose loss the whole can survive) is the first step in salvaging meaning and quiddity for investigation and speculation. Again, to have an integral whole is to have some cluster of properties persist long enough to individuate it qua object. None of these said properties are essential to any given whole *per se*, and it is in their persistence that we consider the dynamic 'hive' of properties and other objects (who also have properties) not as integrated, but as sufficiently similar across and between observations. The 'hive', with roughly-defined borders, structures, micro-agency, properties, parts, events, repetitive features, and perhaps an illusory autonomous quality, manifests itself—and, in fact, *emerges*—in the form of the quidditious whole *simpliciter*. I will now adumbrate Chapter IV's objective: our options for constructing one.

3.4.1. Path via parts

Macroscopic wholes, like mountains and solar systems, are easier to individuate and ascribe properties to, than microscopic wholes, like atoms, cells, and organisms. These macroscopic objects also have, at least in common parlance, comparatively more discernible and distinguishable parts—some of which are often referred to as *proper* parts. The core of the planet Earth is often demarcated as being a proper part of the planetary sphere: it is a geometrically and physically necessary component of the planet, measureable to a fairly precise degree, and without it, there would be no earth. The Earth, with it, is a unified entity. Without it, the earth is non-unified, and therefore not a whole. Proper parts are emphasised as being distinctive and able to be separated and pulled apart, but at the cost of negating the existence of the system qua whole. The human heart is often considered as a proper part, because a) it is seemingly identifiable as a single unit of analysis, and b) it is extremely difficult to conceive of a case where a human could survive without it.

Does the fact that the earth has something that it could not exist without, make it a whole? Is the earth whole because it contains a proper, essential part? The same question can be asked of the human heart: is the human body whole because without a particular, proper, essential part, it could not exist? First of all, proper parts are necessarily individuated, and, depending on the whole, possibly separable. Ultimately, a proper part is a whole within a whole, because it is itself an integrated object. An essential proper part ceases to be a part when it is removed, because the whole could not survive without it. If the whole does not survive, the part no longer has the function of *part*. But the idea of a proper part is the idea of a unified entity contained within, and serving an essential purpose for, the whole. There can be no metaphysical claim for a proper part *per se* without an essence and subsequent quiddity.

The question of whether these macroscopic wholes could survive the loss of such an important piece presupposes the awareness of not only the whole's constituency but that of the part itself. If I were to claim that a human body could not exist without the heart, modern medicine would generally agree. Metaphysical analysis, though, would require us to identify that which is a heart; as such, we would find it difficult to individuate the heart qua part from the body qua whole. In the bizarre scenario where a person's heart had 'been removed' from the body, we would need to clarify that when we say 'the heart' we actually mean the reddish organ with two disconnected and severed primary blood vessels (amongst other debris in and around it and them). It is true that the heart's functions are a necessary sub-function of the circulatory system of any animal's body, but since it was never clear through a refined physicalist and empiricist lens what the heart is, with its functional and systemic interdependence, we cannot claim that the heart is a proper part—let alone an essential one.

Are there ever any real parts? This question is at the heart of any mereology. Most of the authors we have discussed have very clear answers: Husserl's dependent and non-independent taxonomy is one helpful way to deal with *unfounded contents*, for example (e.g. the microphysical volatility we have been discussing).The present re-conceptualisation of emergent wholes includes a nominalist whole ascription, and thus it follows that there can be parts in relation to our quidditious whole *simpliciter*. Since we nominate the whole vis-à-vis the unit of analysis, the parts are left to be determined by the observer and her refined physicalistic and empirical explanation, as language will continue using instrumental and pragmatic simplification. The properties of the whole, on the other hand, whilst far from being observation-independent, are harder to subjectivise. I will provide a clearer answer to this question after I have discussed properties.

3.4.2. Path via properties

There are properties of the whole to consider, as well as properties of the parts. The nature and individuation of the parts, as I have just indicated, seem to be predicated upon the nominal, parsimonious ascription of whole—that which we have been calling MCD. Again, walking the MCD tightrope can easily lead to essential properties (and possibly even MCP), so we need to be careful. The analysis available from *m*CD ascription seems to be the only empirically and metaphysically-sound method, but it can unfortunately result in simply a delayed version of the same hasty whole-ascription (where we say that everything *just is* atoms or sub-atomic particles) that we incurred on ourselves with MCD. This is to say, if we are not precise enough with *m*CD, we end up claiming that the atoms and/or sub-atomic particles themselves are metaphysically unified wholes—which is of course precisely what we have been trying to avoid. If we take the empirically-safer route and put stock in *m*CP, where explanation necessarily looks upward into higher levels, we do not necessarily have a whole, because agency is still dependent on units of analysis. With *m*CP, there would not necessarily and objectively be any quiddity if there was no agency, and vice-versa. Are the properties of objects the only currency in identifying quiddities—and, by extension, the only currency for individuating the whole *simpliciter*?

System-level properties (SLPs) have long been the default evidence for strong (and for some, ontological) emergence. They refer to some configurational, seemingly holistic, adaptive trait of a compound (often carbon-based) unit of analysis. The danger, as we know from Chapter I, is that one might begin to lean towards closed systems, which can collapse into essential wholes. SLPs, though, when placed in a responsible, physicalistic, nominalistic analysis, may prove to be the key to revealing the real, metaphysically-innocent potential of emergence. Within this latter setting, let us look at two different ways of understanding them.

3.4.2.1. Strongly (outright) disassociated system-level properties

Many claims in the metaphysics of strong emergence rest on SLPs that appear to have no causal or constitutive connection—apparent disassociation—with both the parts of the whole as well as the properties of the parts of the whole. In turn they are used as evidence of either some new kind of MCP or DC entity, or some new (kind of) property—despite the refutation or at least contestation from those who question a) the understanding of the whole and its parts, as well as b) the presumption of the whole's integration. Now, is there a way for us to use SLPs for our redesigned purposes? There may be a way to have different categories of agency in our ontologically-innocent framework, because it is possible that, once we have reimagined the quidditious integral whole, we can be flexible with degrees of SLPs. I am slightly worried, though, about how easily one might be tempted to hastily assign unfair agency to one particular kind of system (such as the human brain)—and by extension instantiate what are *prima facie* MCP integral wholes—when *m*CP explanations are seemingly inadequate despite (seemingly) exhaustive investigation. Human cognition, again, is the perennial example of this: we have an increasingly vast knowledge of how the brain and cognition work, yet we still cannot explain mental event E from synaptic pattern P's properties. Those with their novelty-guns cocked will leap at the opportunity for an MCP threshold, and not just as a unit of analysis in a specific investigation. So, for every incomplete explanation there will be a desire for—or at least some attention paid to—autonomous or higher-level agency, whether or not the investigation has been contemporaneously exhausted.

From Chapter I, Bedau's two features of (weak) emergence were:

> (1) Emergent phenomena are somehow constituted by, and generated from, underlying processes.
>
> (2) Emergent phenomena are somehow autonomous from underlying processes. (1997, 375)

The implication of (1) is the rejection of essential wholes and their essential properties in any of the guises for which they have been argued. It would then follow from this that wholes, then, are exclusively composed of underlying objects, events and properties. The suggestion of (2), though, is that we could somehow combine *m*CP, MCP, *m*CD, and MCD—and this is entirely because of the strongly disassociated SLPs. So, counter-intuitively, in these cases, just like everything from autocatalysis to purple cloths, a whole is nothing but its components yet in some way is able to manifest its own quiddity. This might be a better way of describing strong emergence, as it appears to be moving closer to an empirically-safer notion of whole. Quasi-legitimate candidates for strong emergence, then, with their strongly disassociated SLPs, will most likely always be relevant to metaphysics, natural investigation and philosophy in general, because there will most likely always be units of analysis with seemingly disassociated features. It would nevertheless remain possible that that a given strong SLP was a special case of a given weak SLP, or that the former could eventually be reduced to or at least be explained by a set of lower-level properties and approximated conditions.[184] Investigation cannot assume this to be a future impossibility.

[184] Or, we do not know yet and/or have not yet considered how the observer has affected the unit of analysis—and its constituent objects, events and properties.

3.4.2.2. Weakly (nominally) disassociated system-level properties

Weak emergence largely concerns itself with epistemology, as noted in Chapter I. Many things in the world vaguely have features that are:

- ❖ (A1) *seemingly unpredictable from the known behaviour of their constituents; or*
- ❖ (B1) *an unusual global impression of well-explained, patterned local interactions*

In other words, based on what we roughly know now, system S has macro-level features that are either:

- ❖ (A2) *not directly explained by the individual or aggregate behaviour of its lower levels, or*
- ❖ (B2) *largely explained via mCP and/or mCD manifesting a somehow remarkable and unified effect in the mind of the human and the human-programmed observer*

We see examples of A-systems everywhere. In fact every object of perception could in some sense share or at least at one time did share the features of A—especially if we were to consider natural investigation before the time of, for example, Galileo, Newton, and perhaps most relevant to human observation, Darwin. What are some that linger today? The human mind and the atomic nucleus are two of the most obstinate demonstrations of A, though the list is still endless. I suppose that many emergentists and non-emergentists alike would consider there to be a great many kinds of systems today sharing the features of B. A helpful example may be that of a common soap bubble.

There is a need, as with all investigation, to consider the bubble J as a singular unit of analysis, and the primary reason for this is that we can see a difference between J and many non-J's—though this is hardly satisfactory, as we will see. We know that its spherical shape is due to the roughly-even, stochastic distribution of air molecules moving out from the centre mass all confined by the bonded soap and water molecules on the 'exterior'. Putting aside issues of ontological dependence,[185] we would not say J's spherical shape is an SLP type A. We would probably lean more to the type-B side,

[185] These issues are profoundly relevant and painfully inseparable from SLP ascription, and include, but by no means are limited to, the existence of air—matter—'inside' and 'around' J.

because we know generally why it is spherical. Intriguingly, the shape is not entirely determined by the outward push of the stochastic distribution with a molecular net around them: it is also determined by the balance of pressure between that which is 'inside' B and that which is 'outside of it. The 'walls' of B can only confine the molecules within a certain range of pressure on both 'sides'. The mean pressures on either side are not closed off from one another, but are conducted through the molecules, atoms and sub-atomic particles that are transmitting matter and energy between them.

The lessons here are, initially, that this is obviously yet another 'unit of analysis' that is an open system. Not one observer can claim physical compositional and causal closure and therefore metaphysical unity and endurance. Secondly, the recognition of J by the human observer is entirely dependent on the observer's recognition of the soapy film encapsulating the contents of J. J could not exist without its spherical, stochastic configuration, because of the strength of the soap-water molecule capsule. Balloons, for example, can take a number of different shapes because of the strength of the confining material, such as a rubber-plastic compound. Soap bubbles like J appear because they have accommodated 'evenly' distributed air molecules inside a cluster of soap molecules. System J's SLPs, which include apparent duration, colour, surface reflection, and movement in space over said duration, are all derived from the primary SLP, its shape. Shape is an emergent property, then—or, in keeping with our terminology, a weak SLP—in a very weak sense only. J's shape is that which manifests J for the observer, despite the fact that J's shape—an alleged SLP—is directly determined by the micro-level objects, events and properties contained within J. Furthermore, the persistence of J's shape is the only means by which we consider J an integrated, individuated unit of analysis: an unobserved J cannot be a unit of analysis.

The primary and overwhelmingly most important SLP is only such in an incredibly weak sense. The human observer just happens to be able to detect J—and study it—because the molecules and pressures just happened to have aligned in a certain way. Thus what we realise here is that the unit of analysis is individuated by a weakly emergent property, which itself is instantiated by numerous detectable and undetectable *m*CP effects—all operating within suitable ranges of conditions relating to other objects, events, and properties over the duration of their existence. I will now unpack this.

J is individuated (at least in a parsimonious, albeit definitively nominal sense), which means it is detectable (somehow), distinct from non-J (whatever cannot possibly be J, e.g. the sun, a person in another city, even the room in which it is found, etc.), and is in some sense to be considered as a whole (with higher and lower levels, properties, and/or parts). We are able to distinguish J from non-J via properties, which then may manifest themselves as SLPs and therefore contribute to its quiddity. First, with regards to properties, some of them relate to the system-level, which means they are apparent in all aspects of the whole. Some of them relate only to some part or level lower than the system level, and as such, do not relate to all aspects of the system as a whole. Those that relate to the system-level, which are eventually realised as being aggregate *m*CP properties, are erroneously detected by the observer as being properties unique to the system-level. The sufficient persistence of these alleged SLPs over the duration of J's existence give the impression of a self-sustaining system via these higher-level properties (at least some of which are considered by some to be autonomous features—yes, an autonomous bubble!). It is a deceptively apparent co-ordination, fuelled by persistence. For any human observer, a range of conditions is less important for individuation than a range of properties. As such, SLPs are collectively ascribed to J (especially when multiple soap bubbles are observed), which are then encapsulated into a definition of 'soap bubble', which then acts as part of our quiddity.

It seems that we have made, as we indeed often leap at the chance to make, a system, out of illusory higher-level properties—even at a given instantiation and not just over the whole *simpliciter*'s duration. In cases where strongly disassociated SLPs are ascribed, many have granted emergent autonomy to the system qua individuated, integrated whole. In cases where weakly disassociated SLPs are ascribed, many have granted seemingly unpredictable and irreconcilable emergent qualities to the system in order to establish, or at least to contribute to, its own unique quiddity. With both, emergence is the recognition of SLPs and of a quiddity-generating condition of some system. A new vision of emergence may therefore suggest it is the means by which we have objects *per se*: it is the temporary, yet sufficiently-persisting collaboration of objects, events and properties manifesting themselves as quidditious wholes *simpliciter*. The next chapter will consider this issue of a sufficient duration of such a collaboration vis-à-vis system-whole individuation. Before, however, I will look at quiddity-generation via porous emergence.

3.4.3. The porous whole *simpliciter*

With all of the examples above, there has been a recognition of the invaluable and inseparable role of environmental forces—or to be more precise, the *ranges* of conditions—in nominalistic whole ascription. The reaction to this analysis may very well include a critical response about infinite regression and the perceived absurdity of virtually limitless property reduction down into the mysterious levels of the quantum world. As I have mentioned, however, if we want to make claims about structure, agency, autonomous wholes, and genuine novelty, we enter into the metaphysical. As such, we are looking at all possible physicalist descriptions and explanations—even if they 'emerge' from levels below those with which we feel some familiarity. The relevance of environmental conditions and, by extension, a whole's porosity, cannot be understated, and neither can it be ignored. Common objects V and F's very constitution and therefore existence, as with all physical objects, are necessarily predicated upon suitable local conditions. I hope this point has become clear by now, because it will never cease to be outside, or on the perimeter of, any physicalistic and refined empirical metaphysics.

I have tried to slowly reveal a means of salvaging a quiddity *of some kind* in the face of nominalistic whole ascription. This has been made even more difficult with the realisation of an object's dependence on, and porosity in relation to, local conditions. It is unlikely that weak SLPs will ever cease to play a role in any investigation, because investigation is usually prompted by the acknowledgment of that which we consider to be unpredictable or otherwise disassociated higher-level features. The unit of analysis is 'doing something it should not be doing' or is appearing to have characteristics not otherwise known to result from what should be the expected combination of its constituents. The whole, then, can be considered *prima facie* as more than its parts' addition. As we have learned, though, the novelty associated or attributed to the whole was hastily ascribed because of the fractured and/or incomplete understanding of not only the whole's constituents but also the necessary conditions for its existence and persistence. I have tried to show that in every case for novelty, and especially those claiming autonomy, gaps in understanding, whether via selectivity in level of analysis or ignorance of local microphysical agency, do not necessarily justify such novelty or autonomy.

What we know about the whole is not necessarily that which the whole is, even in nominal terms. As such, any purported whole is best identified as a whole *simpliciter*: what we can currently gather about a loose collection of persisting objects, events, and properties, all within ranges of conditions. The nature of this persistence is clouded by multiple, though ultimately incongruent observations with an incomplete catalogue of properties—and this is all exacerbated by a profound difficulty in making metaphysical distinctions between an object and its environment. Thus all investigation, which requires units of analysis, is bound by loose and possibly arbitrary whole, system, part, property, structure and agency ascription. In other words, the physical world, through the lens of a refined physicalism and empiricism, does not present us with absolute units of analysis and essential wholes from which we gather immutable facts. We instantiate, create, and manipulate these units of analysis so that we may understand how the world works. All units of analysis are necessarily porous wholes *simpliciter*. The quiddity that we inject into them largely emerges from system-level properties whose microphysical collaboration we have momentarily and indivisibly acknowledged.[186]

[186] The reader is encouraged to examine the concept of *second-order observation* in social structure and agency studies, particularly the work of Niklas Luhmann (e.g. 1993). This approach helps the investigator do more than simply deconstruct the components of our elaborate social theatre (even such as a Debordian *spectacle*, 1967), but more importantly to reflect on how we observe and to see how our observing processes are fundamentally (and perhaps necessarily) restricted and even invasive.

Chapter Summary

From the first two chapters, we know that the metaphysics of strong emergence needs wholes, and that wholes need quiddities. Wholes are analysable under different, though not necessarily incompatible lenses in two general directions of compositional determinacy and causal powers: macro-level composition (1: MCD) and causal powers (2: MCP), macro-level causal powers featuring downward, reflexive causation (3: DC), micro-level composition (4: *m*CD), and micro-level causal capacity (5: *m*CP). Compositional determinacy relates to how we have come about having a whole to study, whilst causal powers relate to the abilities of any given unit of analysis to instantiate new properties in the world.

Reflexive DC is incompatible with a refined physicalism and empiricism, and MCP seems to necessarily find itself making assumptions of integrated agency. MCD also assumes (at least sufficient) integration to justify individuation, despite being observation-dependent yet requiring endurance. *m*CD simply uses smaller bits on lower levels than those of MCD, so it suffers from a similar incompleteness of object qua whole. *m*CP does not need smaller bits on specific lower levels, as it only requires a directionality. Considering the misapprehensions about the persistence of these smaller bits, it suffers in its application via discrete and independent relations, objects, events, and properties. *m*CP also offers nothing prima facie for structure-agency, because agency is presumably necessarily individuated and distinct from structure.

From Chapter II, all physical wholes seem to be exclusively integral wholes, because no object qua whole maintains their physical parts and properties through multiple instantiations—none of which can be categorised as more or less important than another due to the interdependence through, around and between systems and their components. As such, an integral whole is really just a whole *simpliciter*: a loose but useful approximation of a whole or system-whole. We can individuate a whole *simpliciter* for the sake of investigation via system-level properties, which themselves are never completely disassociated from the whole *simpliciter*. The whole *simpliciter*'s system-level properties seem to be momentary, collaborative conveniences for generating quiddity, granting us access to the wholes as units of analysis. The consideration of local microphysical forces and ranges of conditions reveals an inherent porosity in any unit of analysis, complicating what was already an instrumentally-constructed and therefore fragmented and non-essential quiddity.

As has been admitted, investigation, analysis, description and explanation all require units of analysis. The core problem of the whole *simpliciter* is its persistence and what collection of parts, objects, events, and properties is necessary for such a claim, and this is the focus of the next chapter. We may find that emergence could very well be the quiddity-generating means of individuation in the face of inevitable level-excluding integration.

Chapter IV | How can a whole change?

> All bodies, as they consist of innumerable parts that may be disjoined from them by a great variety of causes, are subject to continual changes of their substance, increasing, diminishing, changing insensibly. When such alterations are gradual, because language could not afford a different name for every different state of such a changeable being, it retains the same name, and is considered as the same thing. Thus we say of an old regiment that it did such a thing a century ago, though there now is not a man alive who then belonged to it. We say a tree is the same in the seed-bed and in the forest. A ship of war, which has successively changed her anchors, her tackle, her sails, her masts, her planks, and her timbers, while she keeps the same name is the same.
>
> *Thomas Reid* [187]

Claims in the metaphysics of strong emergence are claims about dependent novelty. Novel properties are ascribed to an object and singular unit of analysis, often resulting in the identification of an incongruity between lower and higher levels within that unit. Once again, by this reasoning, talk of emergence is therefore talk of wholes—where 'whole' can mean either 'that from which *nothing* is lacking' as well as 'that from which *no part* is lacking'. A mereologically-conscious metaphysics of strong emergence, I have argued, is not afforded the luxury of excluding either of these interpretations of whole (and their related ontological-epistemological conflation), because in order to exhaustively scrutinise the core claims of strong emergence, the non-*ontologically*-reductionist physicalist and refined empiricist is required to consider the nature of the object's integration and subsequent individuation.[188] Most claims made in the metaphysics of strong emergence assume, or ignore the questions of, both integration

[187] *On The Intellectual Powers of Man*, Essay III, Chapter 14, 345 in Hamilton, E (1854)

[188] In other words, the designation of *part* is not exempt from compositional and causal analysis: just because a given table has four legs and a flat surface does not mean that the table is exclusively the same thing as its parts, especially when the composition of said parts is itself begging the question. Entries into the category 'whole' are not untouchable by compositional analysis. That which is 'missing no parts' is equally tenuous as that which is 'missing nothing', and this is, once again, for two reasons. (1) We are using a refined empirically-conscious physicalism, which cannot exclude any physical level of analysis. (2) That which is 'missing no parts' (which pleads for super-semantic exemption) depends on essence (or some unifying supersummative element) or at least quiddity to, at best, identify inclusive components, and at worst, identify exclusive levels of analysis. The totality of X can traditionally only be known when we assume what-ness. I am trying to show how we can have quiddity without a claim for totality.

and individuation, and therefore, I believe, overlook the necessary transformation of the whole *per se* to the whole *simpliciter*: the unit of analysis to which we ascribe weak system-level properties and nominally-discrete parts. This device is the first mereological step in the realisation of the instrumentality of any ascription of quiddity—or what-ness—via convention or definition (and not essence).

Investigation[189] and any empirically-conscious analysis of the claims in the metaphysics of emergence need individuated objects, but justifying individuation requires integration—which in turn can be known aprioristically from the classical perspective or via quiddity (which can be known both via essence *and* convention). I have taken the path of the conventional (which includes definitional) quiddity: as such, the construction of an object's what-ness is, in this mode, potentially a linguistic-historical issue, and not one of logic or essential truths. Simply because the object's system-level properties have been ascribed to that which is physically non-integrated, ascribing a quiddity to a given object requires assumptions. I have laid out many of those assumptions in this thesis, and there is one more area left to treat. As an object inevitably faces relentless compositional and property alteration over any micro-duration, on what grounds can we, with a full awareness of the instrumentality of such an integrating and individuating assumption, justify any particular quiddity ascription? Hence, this chapter is about how the whole *simpliciter* and its potentially emergent system-level properties could be understood as surviving change.

4.1. Identity as preceding persistence

The problem of change is even older than the problem of 'the one and the many'. Heraclitus, for example, believed that nothing could survive the replacement of its parts, and, thus, no person could step in the same river twice. The refined physicalist[190]

[189] Investigation, as I have been using the word, comes to mean, simply, 'learning about the world.' This would accommodate observation, description, explanation, and in some cases, future estimation (sometimes called prediction). As such, both the natural and social sciences converge on a (consilient) shared purpose.
[190] Again: I am still using this word in the *not-necessarily-ontologically-reductionist* variety: it simply means that we grant ontological priority to that which is recognised by physics—and this does not require or imply ontological primitives (as standard models are incomplete and do not necessarily use

and empiricist would tend to agree about the river, because there is an unfathomably low probability of the recurrence of the exact same molecular (to pick one level) constituency at both t_1 and t_2.[191] The stunning irony in the former's claim is not that he was right about the molecular volatility of any volume of liquid water under standard earthbound conditions. It lies in the fact that his own reasoning for challenging the *endurance* of 'the river' also applies to the persistence of the person stepping into it. With ninety-eight percent of an average human body 'rebuilding itself' in about three months—with pervasive and instantaneous shuffling of matter and force throughout the porous body—the person is truly *analogous* to the river. The same person cannot step into the same river twice because there is no same river and there is no same person—with a physicalist presupposition and without any non-physical criteria for personal identity.[192] [193]

As was the emphasis in previous chapters, because we do not have a complete set of properties, and are therefore unaware of what the 'river' actually is in physical terms, we can never know whether the same 'river' is stepped in at both t_1 and t_2. Persistence is inherently problematic in a refined physicalist ontology. More importantly though: even if we had a whole, how could we make claims in the metaphysics of strong emergence if the whole had not persisted, or the whole was not the same?

---

or care about primitives anyway). Though I prefer the term *metaphysical naturalist*, I have continued using 'physicalist' in the hopes that its ontologically-reductionist connotation will fade away gradually.

[191] And, we do not need to pick the molecular level—despite the significance of an H_2O molecule being commonly called 'water', which 'constitutes' rivers. As we have seen in the previous chapter, there is profound change at the sub-molecular level happening at very small time-scales. Another key feature of this example is the consideration of the volatility of any volume of water: at any given moment, untraceable photons and various forces and forms of energy are stirring up the water molecules in a given container—some of which may be released from its liquid state and vaporised. As such, the volume of liquid water—and therefore the physical constituency—has changed with a single vapour.

[192] Referring back to Chapter I, I remind the reader of the implications of rejecting a physicalist ontology: to discard all physics and any of the special sciences on the grounds that they cannot inform the description and explanation of any particular higher level of analysis.

[193] This present work has deliberately avoided issues of personal identity and humans-as-wholes, so I will not continue to explore this issue. But, it should be recognised that as we continue down this refined physicalist path, this is most likely inevitable.

4.1.1. Leibniz' Law(s) (LL)

As referenced in Chapter II, Leibnizian calculus tried to bridge quantitative and qualitative identity—both being necessary preconditions for persistence.[194] One must know 'what' and 'which' object is persisting in order to determine if it does indeed persist. There are, however, two varieties of it, both found in the Clarke (GM, GP) correspondence.

> LL_1: (indiscernibility of identicals): For any x and any y, if x = y, then x and y have the same properties.

Many philosophers and logicians, and most mathematicians accept this is as plainly true and uncontroversial. Unfortunately, it is limited to the mathematical realm. If I purchase two sealed copies of the same studio album from the same retailer, I can give either one as a gift. They both belong to a set, series, or category, of which all members are equal in their membership and possessing only numerical difference. This membership is not metaphysically significant, though, as I may create any set, series, or category at whim allowing me to organise any group of objects. It may be epistemologically significant, but that is outside the scope of this section. Where it becomes fuzzy is in its ontological status; but since we have been using a physicalist ontology, it does not carry much weight there, either. Numerical identity must operate under the assumption that the objects are qualitatively identical, and so properties are unequivocally part of the claim. Qualitatively, there is infinite difference between the albums: the condition of the plastic packaging, the nature, location, and distribution of the chemicals inside and outside the package, and the microphysical variation in environmental conditions in, through, and around each package, to name just a few arbitrary groupings of differences. Assuming both are spatiotemporal wholes, and even if we grant that there is an 'x-whole', there can be no assimilation between it and

[194] It should be self-evident, I would think, that identity must always be dependent on, and therefore methodologically precede, persistence, because any instantiated object needs no identity if it is not subsequently instantiated. The only objects that I can think of that are not subsequently instantiated—and yet are considered objects in some sense—are the particle/antiparticle mutual annihilation within quantum foam (though technically an unobserved phenomena). Yet, the duration of these particles (imagine tiny bubbles popping out of a soapy foam), even if they were measured, would still be greater than zero in some time measurement. This means that the duration could be cut in half, and the objects could be seen to be instantiated at two points (implying persistence of a given object), and therefore identity would still be preceding persistence.

a ‘y-whole’. Their mutual membership in a set is all that connects them, and that is not enough to equate them.

It is rather difficult to imagine any physical object being isolated (at any level of analysis) or be deemed a closed system (at any level of analysis), and as such, could ever be considered as comprehensively-physically identical to something else. This notion for me in our present analysis is difficult because it assumes a complete set of properties, which itself in the claim evidently persists. So, in this single claim, even today seen by many as essential to mathematics (if not often in other disciplines), we have one possible source for the modern misapprehensions about wholes and persisting objects. I will return to this later in the chapter.

> LL_2: (identity of indiscernibles): For any x and y, if x and y have all the same properties, then x is identical to y.

The inverse is even more problematic, and for reasons seemingly unaddressed by the existing literature. Max Black’s over-celebrated critique (1952) uses a parallel or symmetric universe approach to show that two spheres can be numerically different whilst being qualitatively identical. It is not clear why this critique has found itself included in metaphysics textbooks, as it really makes no contribution to the debate on identity and change at all. (1) Parallel universes are of no interest to the refined physicalist and empiricist, as they presently cannot be observed or measured, and remain mathematically-theoretical. (2) If there are two spheres (sharing at least some of the same properties), they must not occupy the same space (assuming standard earth-bound conditions, as the author is doing). If two objects are not sharing the same space, they cannot be subject to the same precise microphysical conditions—even if such conditions could be exhaustively measured. So, Black goes about criticising LL_2 in a bizarre and irrelevant way.

What is far more interesting—and what has consistently been a target of this present work—is the brazen assumption that two objects might actually have all the same properties.[195] This is to say that $\{p_1 \ldots p_n\}_{P1} = \{p_1 \ldots p_n\}_{P2}$. In other words, one is claiming

[195] This is one example of contemporary philosophy utilising devices (like Black’s spheres) that are absurd, unobserved, and presumably non-existent. On this point on properties: (1) Person A claims that two objects have the same properties; (2) Person B says that such a situation could not happen; (3)

(1) to know the complete catalogue of properties of one object, (2) to know the complete catalogue of properties of a second object, and (3) to verify that both sets are identical and enduring. (1) and (2) have already been dismissed on numerous occasions. As mentioned in Chapter III, to claim a *complete* catalogue of properties is to claim an *enduring* set of properties.[196] Observations are not precisely the same as instantiations, and property ascription on any level of analysis is necessarily made by a spatiotemporal object—whether it is a human brain or a supercomputer at CERN. Instantiations are used to show the *intervals* of our observations, and not the observations themselves.

A person tells me they have placed a piece of fruit on the table behind me. There is a significant duration of time between the gradual turning around in my chair, the receipt of the reflected red light off of the fruit, and the image recognition of the fruit as an apple. I will most likely see first that it is red, and then briefly after that I will notice things about its shape that will have some degree of similarity with my idea and mental image of an apple. I may even consider very briefly that it is an apple simply because it is red, before I have had adequate time to inspect it or at least match it more carefully to my idea of an apple. There may not necessarily be this unidirectional sequence, but there most certainly is a duration of this property ascription, and it is greater than zero in any unit of measurement.

The point here is that observation-dependence implies spatiotemporal observation, and property ascription is itself a dynamic and diachronic process. A photographic snapshot is the closest analogy of what an instantiation might be, and the properties that we gather from one will never be a complete set. Despite this inadequacy and obvious bias, this is generally how we understand not only objects, but properties, and

Person A says, "ok, but imagine that they did." The point here is that these thought experiments should not be used for metaphysical analysis, because the example can have no correlation to real-world objects—and thus we could have no way of showing that any analysis is useful or correct. If I were to devise a thought experiment about parallel universes, and that all but one of the properties of Earth-2 were identical to Earth-1, what possible function could such an experiment play? How could these parameters be taken seriously in contemporary academic metaphysics?

[196] This challenge against enduring sets of properties is, of course, on top of the challenge against the possibility of a complete set of properties (discussed in Chapter III). In other words, even if we were to grant that a human observer had access to the complete catalogue of properties of object X, we would still need to show that this catalogue has persisted (or where all properties of X were instantiated at t_2 as well as t_1).

our language reflects this tendency. These statements from Leibniz are helpful, and they are practical, but they do not reflect the natural world. In other words, they seem incompatible with a refined physicalism and empiricism, because they do not coincide with our coded sensory data. Furthermore, it seems that they continue to be major obstacles to contemporary philosophy—particularly in the metaphysics of strong emergence.

4.1.2. Conditions of change

For an object to change from t_1 to t_2, it be must be both the same and a different object—and this has always been the source of the problem going back to the Eleatics. An unripe banana at t_1 must be both different from, and the same as, a ripe banana at t_2. As I have been trying to illustrate throughout this work, any whole or unit of analysis necessarily experiences at least microphysical change from t_1 to t_2; as such, for a whole to persist, and possibly even manifest emergent properties, it must change whilst staying the same. This is the crux of the issue, and this where emergence, wholes, and persistence most vividly come together.

Following from Heller (1992), Merricks (1999), Haslanger (2003), and Wasserman (2006), three conditions expressed in similar ways have been raised that demonstrate the fundamental contradictions that the idea of change presents us with.

- (D) *Difference condition: the unripe banana is green and the ripe banana is yellow.*
- (I) *Incompatibility condition: nothing is both green and yellow.*
- (S) *Sameness condition: the unripe banana is the same as the ripe banana.*

The (D) argument is that change requires a difference in properties, whilst (I) claims that there must be at least two properties that are mutually incompatible. (S) implies that both properties must attach themselves to something shared by both B_1 and B_2—which means that this 'something' must persist *in some way*. First, we have a vague idea of what 'ripe' means—as it reduces to edibility and the balancing of acids and sugars in the fruit. But in this example, we are meant to think that 'unripe banana' is synonymous with 'green banana'. If we are to be precise—and this should be the underlying aim of metaphysics—there is no such thing as a metaphysically-significant categorisation of 'unripe banana'. There is more than just 'ripe' and 'unripe', for there are degrees of ripeness and edibility. But, since we are really just focusing on the green-yellow distinction, we can turn to optics.

As has been mentioned earlier in this present work, colour-as-property ascription is fundamentally flawed and ultimately a linguistically-instrumental tendency—and this is because of our knowledge of optical physics. A macroscopic object like a banana may seem to have a predominant 'colour'—the most perfect, prize-winning banana will probably feature a seemingly uniform sunny yellow hue—but its *apparent* uniform light reflection does not require its *actual* uniform light reflection. There is still a vast, indeterminate amount of light reflected that would not be considered yellow that is simply outmatched by what we are calling the reflected yellow light.[197] Furthermore, a banana may be considered still ripe despite it having a few tiny brown spots or a greenish-yellow tinge near the stem—we would still call it 'yellow' as well as 'ripe'. For one banana to be considered as green and another as yellow—and to use this as a basis for a defence of metaphysical change—is folly. There is a gradual, dynamic process of fluctuating acid and sugar levels that is directly affected by the gradual, dynamic process of fluctuating levels of electromagnetic forces with relatively stable (though not static) temperature and pressure conditions. Picking a moment (t_1) and labelling it 'green' (and therefore 'unripe') is useful for farmers and linguists, but offers no precision for metaphysics. There is no green and yellow distinction—there is simply a vague difference in properties between multiple instantiations.

The claim of (I) is immediately thrown out as a result of all this, as something can indeed be both green and yellow—in part because our conception and ascription of either was already incomplete and arbitrary. The real challenge is that of (S): what do both have in common? What holds it all together? What persists? Ultimately, these conditions above all assume that $B_1 \neq B_2$, where B is the banana, B_1 is the unripe banana, and B_2 is the ripe banana. I write ($B_1 \neq B_2$) because B is seen as a particular object, that persists, but with variations in its properties. Yet, it requires a persisting

[197] There is also a possibility here of the mind wanting to see the yellow only—or of deliberately or subconsciously ignoring/overlooking the small patches of brown or the greenish yellow tinge near the stem, for example.

particular, B, to which those different properties are ascribed. Since $\{P_1 \ldots P_n\}_{t1} \neq \{P_1 \ldots P_n\}_{t2}$, and a set $\{P_1 \ldots P_n\}$ must relate to an object, the comparison must be relative to the same object (B). Hence, describing that which is B requires a quiddity, which is indicative of persistence. Though B has a different set of properties at t_1 that it has at t_2, we are still calling it B. So, we are back to the problem of our justification for object individuation. If an object's set of properties change from one observation to the next, on what grounds can we say that it is the same object? This is the connection between persistence and quiddity-generation: I believe it is likely that with many common objects, claims for individuation (and subsequent system-level property ascription, all the while assuming physical integration) and claims for persistence are mutually reinforcing. Object O persists because it has been individuated, and O is individuated because it persists. And, when asked to justify either or both, we typically fall back on some kind of quiddity. Since we know that a non-essential quiddity is necessarily constructed, and that a non-ontologically-reductionist physicalism and a refined empiricism cannot recognise essence, it seems to me that the individuation and persistence of common objects is the ultimate collaboration of local observational conditions and some conventional quiddity. From which does Banana B arise, and why should we feel compelled to declare property alteration and potential irreducibility? It could very well be the case that not only are the properties of B epistemologically emergent, but that an integrated and individuated B and its very quiddity possibly are also.[198] I will return to this thought at the end of the current chapter.

4.1.3. The Ship of Theseus[199]

The perennial question of *The Ship of Theseus* (S) has been—even over the last century or so—the centrepiece of many debates on metaphysical change. If one changes the parts of a whole arbitrarily at different times, to the point where the eventual whole is composed entirely of non-original parts, how can one still consider

[198] It might seem as if this statement is hinting at a kind of epistemological emergence of objects: this is about right, though later I will label it 'weak object individuation'.
[199] The puzzle has taken on a number of forms, with subtle differences, including 'Heraclitus' River', 'Plato's Carriage', and 'Locke's Socks', among others.

the whole as *the same whole*? It is ultimately about the nature of persistence, with an explicitly mereological dimension. Most of the solutions[200] offered have preferred the use of parts over properties in ontological terms to identify the ship, though some common treatments use a kind of narrative, epistemological criteria for identity (where the ship is the same, because it has continued to be observed as such over time).[201] This means that S will be analysed purely in terms of its proper parts ($P_1 \ldots P_n$)—where each proper part is one plank of the ship, and the sum of the proper parts is identical to S (if only at t_1). In this way, as the standard analysis goes, as each plank is replaced, one by one, there is a continuity between observations because some planks are present at successive instantiations. Even as they are removed, and replaced by new planks, the existence of those new planks will overlap other successive instantiations. Without the use of properties, but with an arbitrary designation of *part*, it is possible to claim that part P is present at both t_1 and t_2 thereby verifying the object's persistence based on the presence of P at both observations. With the existence of P at both t_1 and t_2, S, as the argument goes, persists, and we still have *The Ship of Theseus*. The narrative strategy works, then, if we exclusively acknowledge proper, discrete parts.

Parts (or, rather, objects qua parts) and properties are the two key types of currency[202] in the validation of a whole's persistence—and, by implication, identity. Claiming parts—as we have seen, especially in the *proper* sense—creates problems equally as challenging as claiming wholes. With Scholastic integral and Leibnizian mixed wholes, there are parts that *are* dependent on the whole for its existence (possibly essential parts) as well as those that *are not* dependent on the whole for its existence. Concerning dependence, certain parts would need to be present for it to be *The Ship of Theseus*, and—what is more important—these parts at t_1 that are required for it to remain *The Ship of Theseus* at t_2. Concerning independence, there are parts that

[200] Notable treatments include Kripke (1980) and Chisholm (1967, 1973, 1976). There are, of course, other extreme treatments involving the 'essence' of the ship or the continuity of its 'owner' (Theseus). These latter cases are of no interest simply because they are incompatible with a physicalist base: one must use as many of the properties at our disposal in informing our description and explanation.
[201] This is much like the quote from Thomas Reid at the beginning of the chapter, which says that, even though we know things are constantly changing, language cannot afford a word for every stage or form of the object changing.
[202] Objects (qua parts) and properties are joined by events and conditions as the four possible currencies in description and explanation. All four can also be reduced to, as well as manifested as, properties. In short, what appears to be a property or an object might also be an event with certain conditions, and vice-versa, but we always come back to properties (as coded sensory data, see Introduction and Chapter I).

come and go before, during, between, and after t_1 and t_2. Since we have no indication of which (if any) proper part is essential or dependent, and since all parts of S evidently belong to the same ontological category, we have no choice but to assume that all P's ($P_1 \dots P_n$) are limited to numerical difference. There must be at least one proper part present at multiple successive instantiations, but we are still dealing with an inconsistent set of parts—with no declaration of essential parts and no categorical distinction between planks. No single plank is more important than another, which means that, of the proper parts included in the analysis (meaning, planks), there is no particular plank needed to identify the ship as S.[203] So, how does S persist if there is nothing linking all the planks together?

Strangely enough, even in the narrative view—though evidently unwittingly so—S persists, not via overlapping parts, but via shared properties instantiated at successive observations. A plank, as a proper part, needs to be individuated, which would require a seemingly-cohesive set of properties. How could we consider a plank to be a unified whole—a unit of analysis—which would be necessary for it to be a proper part?[204] Of the materials contained within an alleged whole, some must be assembled into parts—and so the process of individuating parts is also a process of binding properties together into smaller units serving (what will eventually be seen as actually just the idea of) the whole. At least in common language, P_1 is a discrete object that bangs on the ground when we drop it. It collides and it supports, and it can be felt, eaten or burned. P_1 itself has been crafted, and yet it also seems to be able to 'do' things on its own—like bearing the burden of a man standing on it. All these traits, in addition to its physical ones, point not to a measureable, reproducible, or even universal conception of 'plank' or P. These features all point to an independent conception of S—and by extension, P—to which we connect them. We do not connect the turbulent and volatile microphysical events and conditions happening in, through, and around what we have been calling S or P, and as such, we have not considered these events and conditions

[203] This is all assuming that there was not, for example, a unique masthead that had always been a signature feature of S, that also needed to be repainting, repair, or replacement.

[204] Detecting a persisting plank of common wood, to begin with, requires suitable lighting and atmospheric conditions. The properties needed to individuate one or more planks (isolated or aggregated as *S*) are directed through the five senses: how it looks, how it feels, how it smells, what sounds it makes, and what flavours it may have. If these vague, qualia-oriented properties are more or less available at multiple observations, we would probably declare that the seemingly-cohesive assemblage of properties persists.

as relating to our pre-existing definitions and understandings of S and P (i.e. an 'idea of an object'). We chose our unit of analysis, made it a whole by determining its system-level features, as well as its parts and their relations, and then discarded whatever physical and ideational descriptions and explanations that allegedly were not informing such a conception.[205]

The entire process of part or whole ascription depends on the stretching of microphysical properties across multiple instantiations. This is the case with alleged wholes on all known scales: from the molecule to the mountain, our claims of a) unity and b) persistence require the wilful avoidance and sophisticated refusal of uncomfortable compositional and property alteration.[206] If properties are used to identify S at t_1, then properties must also be used at t_2 in order to claim S's persistence. Needless to say, S will not pass the test of perfect property persistence—what we have been calling endurance—if for no other reason than the replacement plank's wood is new and from a different tree.[207] As was the case with the examples from Chapter III—after having discarded essential wholes—S does not endure because we a) do not have a complete catalogue of properties at a given, single observation, and, by extension, b) cannot claim that a set of properties can be instantiated at multiple, successive observations. It may be said to persist in some way, but with our refined physicalism and empiricism, we have already ruled out endurance.[208]

[205] Philosophers often seem to give weight to 'common sense' or common language, as if learning about the world is not just restricted, but confined by some *zeitgeist*. *The Ship of Theseus* is one of countless examples of philosophers pretending to pitch a so-called conundrum that they themselves invented as being a barrier or challenge to knowledge or understanding. We should know by now the limits of language, and it is my opinion that we should not seek to exacerbate the situation any further.
[206] And, this is precisely what we discovered with the non-essential integral or mixed whole: the Scholastics knew they had to account for parts and properties that did not seem to fit their understanding of common objects.
[207] An endless list of differences would ensue. Besides the obvious mass and weight changes, the shape and way the parts fit together will be slightly different. Microphysical and chemical changes are never irrelevant.
[208] As we have said, distinguishing between the mereological question of the whole as (a) that from which nothing is lacking and (b) that from which no part is lacking, is irrelevant—for many reasons, but especially this issue of persistence. If neither an object qua whole nor an object qua part is physically persisting, claims for system-level properties are inherently tenuous.

4.1.4. Entia successiva

Despite rejecting four-dimensionalism (4D), Chisholm (1976, 98) would consider S as an *ens successivum*: a "logical construction upon the various things that may be said to do duty for it." But, he admits (212) that his *ens successivum* was inspired by notable Christian authors, including Augustine:

> Thus it is always that when any single thing is composed of many parts which do not coexist simultaneously, the whole gives more delight than the parts could ever do perceived separately.
> (*Confessions*, Book IV, Chapter IX)

He gives a further reference to Aquinas, who says that some of [an object's] parts do not coexist with others of its parts: *una pars non est cum alia parte* (*Commentary on The Sentences*, Book I, Distinction 8). Chisholm's formulation does not emphasise the temporal in the same way that these authors do. The reader is reminded that the Scholastic authors, including Aquinas,[209] recognised the need to discuss the *totum integrale* that had inessential as well as essential parts, largely because they had to overcome issues of partial non-persistence. Chisholm is emphasising the unifying force—without going as far as an Aristotelian *form*—that the discontinuous and disconnected parts all "do duty for." In all of these cases above, though, all of the whole's parts cannot coexist: but, that is not the ideal expression. The better phrasing would be: some parts cannot coexist, because they happen to be spread out at different times and perhaps in different places. The other interesting feature of this *ens successivum* is that the whole's parts are discrete, which means that they are distinct from one another, or where $P_1 \neq P_2$. This appears to have no connection with the discrete temporal parts found in most 4D accounts (as we will see in §2.1). However, it threatens to collapse into the exclusive usage of proper parts, which must be discrete.

All that being said, here is how Chisholm would see the *Ship of Theseus* problem, in light of his *ens successivum*. Some planks were joined together at t_1, some of those were removed at t_2, and some of those at t_2 were removed at t_3. The aggregates at t_1,

[209] *On Truth*, Question 25, Article 5 (1954)

t_2 and t_3 are all different, and yet they all have something in common: the aggregates at each observation all perform a service for S at their own times. In this way, the *Ship of Theseus* is a construction of various things doing different things at different times in the service of its own conceptualisation. The consequences for this are rather shocking.

- *(A) S only exists if S persists.*

This does not mean that S must be successively instantiated in order for S to be instantiated. Nor does (A) mean that the constituents of 'that-which-would-become-S' existed before S, which would, of course, be true. It *does* imply that one's acknowledgment of its existence is only possible when the physical constituency of S is at least partially successively instantiated. In other words, observation is required for S to exist, but only via multiple, successive instantiations. There is no major problem here, as we have been using a physicalist-nominalist framework anyway. But, there are conflictions with several other implications.

- *(B) S only persists if S exists.*

This is to say that S must be instantiated in order for it to persist: it must have appeared at t_1 in order to be considered as having persisted from t_1 to t_2. On the surface this is incompatible with (A) because in (A) we seemed to establish the opposite sequence from (B). S can only be known via multiple instantiations; and yet, in order to for S to be recognised as such at any particular instantiation to validate a step in its persistence, it must have some distinguishing feature of S. This may dealt with by once again changing 'S' to 'the physical constituents of S at time-t'. As such, we are saying that the physical constituents of S at t_1 must at least partially be present at t_2 in order to claim persistence of S. It seems plainly obvious, and not much of a digression from our reformulation of (A).

- *(C) S only exists if ($P_1 \ldots P_n$) exist.*

This means that S is only instantiated if it has discernible proper parts. Even if we were to take the most ancient of ships, made *entirely* out of wood, with no miscellaneous non-wooden parts and materials, we would still have constant chemical and microphysical activity. The original question assumes degradation of the planks over a great length of time. No matter how long that duration, the planks are changing chemically: even a single trip out on the water will have altered all of the most basic physical and chemical measurements. Change is inevitable, even when we exclusively analyse discrete, proper parts. The identification and usage of proper parts to claim persistence *of some kind* will always be as parsimonious and instrumental as the unity of the ship itself. Furthermore, a priori parts and 'properness' are untenable as they have no essential system-level integration, unity, or founding relation.

- *(D) ($P_1 \ldots P_n$) only exist if S exists.*
- *(E) ($P_1 \ldots P_n$) only exist if they persist.*

(D) is the main thrust of the original Chisholm quote, and a necessary feature of essential wholes. It is also the inverse of (C), where the presence of proper parts precludes the instantiation of the whole, thus making the whole eligible for persistence. In order for the proper parts to "do duty" for the whole, we would have needed an approximated whole in mind before a) identifying proper parts, b) identifying how they work together, and c) identifying how they work together to some end, purpose, or duty. As we have found from the beginning of this present work, wholes appear—at least in an incomplete sense—before parts, because 'part' can only be applied to a rough idea of 'whole'. Two parts share a relation to that which encapsulates both, but no part has relations exclusively to the whole. An ontologically independent part cannot be a part. Hence, unified wholes could never consist of proper parts—and this is a major problem for mereology in general as admitted by Simons (1986, 290-2). S must exist in order for ($P_1 \ldots P_n$) to exist, so ($P_1 \ldots P_n$) must persist if S is to persist; however, (C) prevents ($P_1 \ldots P_n$) from persisting because of inevitable microphysical changes. (E) contradicts (C) and (D) indirectly, and yet, it is a necessary outcome of the original statement.

- *(F) ($P_1 \ldots P_n$) only persist if they exist.*

Intriguingly, this is not as plain a statement as it may seem on the surface. Even though not all wholes need proper parts, enduring, unified, and/or essential wholes necessarily require them. The whole comes first, then the ascription of parts—in this case proper ones—which are then necessarily brought along in the whole's persistence. But, then we are back to where we started, because the proper parts could not have endured from t_1 to t_2, as stated in (C).

The *ens successivum* is, I believe, a useful device for understanding the problem of *The Ship of Theseus*, despite its self-contradictory implications. When it has been scrutinised, we find that it reveals not just the inherent restriction that our language has on philosophy and investigation, but also the instrumentality of it all. Language concerning change is, at its best, pragmatic, but most commonly, it is deceptive: Chisholm unknowingly shows us that all wholes—and quite conceivably, most *nouns*—are *entia successiva*. An *ens successivum* is an elaborate construction of a persistent quiddity—a necessary ingredient of language and investigation.[210] Appealing as it may be, as I have presented it, it is interesting to note its relationship with the mainstream options on temporal parts (TPs). 4D in particular, which includes perduring options directly opposed to Chisholm's quasi-enduring *ens successivum*, may offer some useful approaches to understanding persistent, emergent wholes.

4.2. Persistence and temporal parts

Persistence is usually divided into two general camps of endurance[211] and perdurance.[212] We have examined endurance and generally found, within a refined physicalism and empiricism, nothing tenable due to the overwhelmingly unlikely

[210] It has a unique relationship with another term used elsewhere by Chisholm and many others: the *entia per alio*. I will return to this relationship towards the end of the chapter. For now, let us continue to consider how the persistence of wholes is dependent on its quiddity, and how that quiddity is itself emergent.

[211] Notable three-dimensionalists include: Baker (2000, 1997); Burke (1994a, 1994b, 1992); Chisholm (1976); Doepke (1982); Gallois (1998); Geach (1972a); Haslanger (1994, 1989a, 1989b, 1985); Hinchliff (1996); Johnston (1987, 1992); Lombard (1994); Lowe (1989, 1988a, 1988b, 1987, 1983a, 1983b); Mellor (1998, 1981); Merricks (1999, 1994a, 1994b); Oderberg (1996, 1993); Rea (2000, 1998, 1995); Simons (1987); Thomson (1998, 1983); van Cleve (1986); van Inwagen (1990a, 1990b, 1981); Wiggins (1980, 1968); Zimmerman (1999, 1998a, 1998b, 1997, 1995).

[212] Notable four-dimensionalists include: Armstrong (1980); Balashov (1999, 2000); Broad (1923: 54–5, 63, 1933: 141–66, esp. 166)); Carnap (1967, sects. 128, 159); Cartwright (1975); Goodman (1951, ch. IV, sect. 1); Hawley (1999); Hudson (1999); Jubien (1993); Le Poidevin (1991); Lewis (1988a, 1986a: 202–4, 1983a (plus postscript B)); Lotze (1887, chs. 1–4); McTaggart (1921: 176–7); Quine (1960, sect. 36, 1963, 1976b, 1981: 10–13); Russell (1914: 112ff., 1927: 243ff., 284–9); Smart (1972, 1963, ch. VII); Whitehead (1920); Williams (1951).

existence of a wholly-persisting object. Perdurance, though, is quite diverse in its 4D applications. David Lewis coined the term in *On The Plurality of Worlds* (1986), with Ted Sider (e.g. 2001) and Mark Heller (e.g. 1990) generally receiving credit for carving out the opposing views of stage and worm perdurance, respectively—which I will discuss in the coming sub-sections. Both of their views, however, involve the use of TPs, something that the endurantists explicitly reject. Considering my apprehension about an ontologically-significant 'part', a temporal part doctrine may prove to be controversial despite its appealing connection to 4D, dynamic, diachronic, and possibly, emergent thinking.

4.2.1. Discrete TPs

Michael C. Rea (2003, 246) and others try to describe 4D in relation to presentism, the doctrine that only present objects exist. 4D, he writes, claims that there are past and future objects equally as real as present objects.[213] We are to see non-present objects in the same way that we see spatially distant objects: "they exist, just not here, where we are." These objects exist in this way because they are constituted by both spatial and temporal parts—where both can be individuated, and in a sense, are discrete. Here are two descriptions of TPs:

> (1) [They are] parts that exist solely because of [the object] existing at a certain time. (Simons & Melia, 2000, 59)

> (2) [Relevant to an object O existing from t_0 to t_3] A temporal part of O, called P, is an object that comes into existence at some time after t_0 but before t_1; it goes out of existence at some time after t_2 but before t_3, all whilst taking up some portion of the space that O takes up for all the time that P exists. (Thomson, 1983, 202)

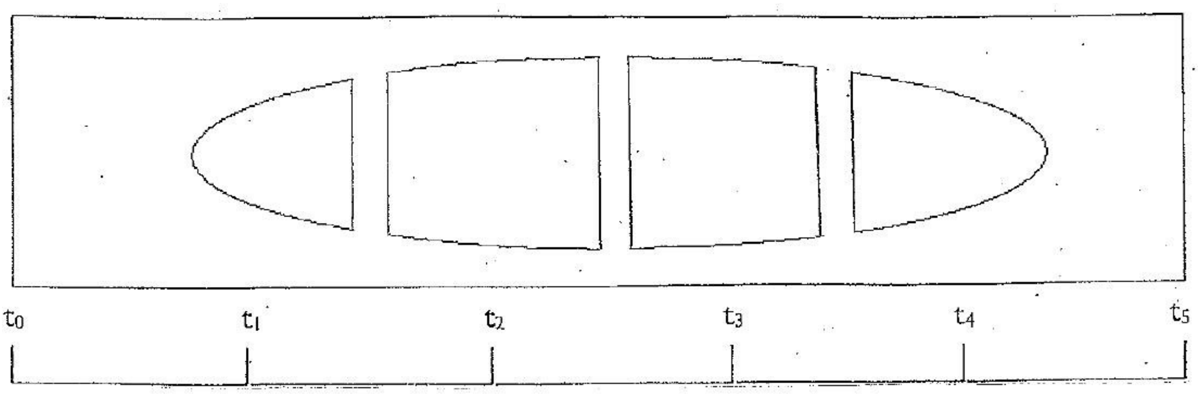
In (1), 'Peter' is constituted by a 'Peter-at-t_1' part, a 'Peter-at-t_2' part, and so on. So, 'Peter' is, but is not limited to, *Peter-five-years-ago* as well as *Peter-five-years-from-now*. As Heller sees it (e.g. 1990, p5 or 1992, p697), it does not matter that there are, evidently, an infinite number of temporal parts, considering the infinite number of

[213] I have no particular interest in this debate, as I see it as a trivial and anthropocentric non-problematique—especially considering the limitations we know we have on our measuring, understanding, and experiencing of that which we call 'time'.

possible instantiations on which we could apply a temporal-part label. As such, *Peter-at-31-October-1995-at-17:50:08:34* is also a temporal part. It would seem strange to consider thinking that TPs could be understood in this absurd, discrete way, but, alas, they still are. Judith Thomson's description in (2) actually refers to a TP as a *spatial object*—which pushes the absurdity—and the Einsteinian abuse—even further. Of course, in her work, Thomson is actually trying to discredit 4D, but in the process, has either misunderstood TPs completely or is revealing them to be inherently flawed as they had been presented in at least one place by at least one author.[214]

Directly related to this *Peter-at-time-t* view, TPs are also to be understood as 'time-slices', where a 'scattered', persisting object is like an oblong-shaped strip, cut vertically. Consider only an x-axis horizontal line, where time flows in one direction. Place the strip on the graph, where one end rests at t_1 and the other end rests at t_4. Now consider the strip as chopped into four similar portions at each instantiation, as shown below in Figure 4.2.1.

Figure 4.2.1.1: The Time-slice View



When all of the slices are considered together, they collectively represent a persisting object. One portion alone is an instantiation of the object (e.g. *Peter-at-t_1*), but when blurred together over many *Peter-at-time-t*'s, we may be justified in claiming that Peter persists in some way. Now, this does seem to make some sense, particularly because it does not necessarily challenge the nominalistic picture—as well as some features of the physicalistic picture. There is, however, a major concern over the implications

[214] She pays explicit attention to Cartwright, R (1975) and Wiggins, D (1980).

of separating the time-slices, as described above, with the matter and force that compose each slice. The time-slice view makes sense, but not by itself, because—and this may be intentional for some proponents of 4D—it is an attempt to isolate one 'dimension' from the other three. Time, as we are learning more and more, is quite possibly best described as an observation-dependent *property* (or 'accident') of matter and force,[215] and as such, it can never be individuated.[216] Therefore, a series of time-slices, like the one above, must always be juxtaposed with a three dimensional space. Once this is done, we would clearly have a difficult time reconciling these instantiations with an arbitrary and ultimately incomplete catalogue of physical properties. The time-slice view, it seems, is a useful way of seeing persistence as, in part, a series of moments or instantiations, which echoes our own inevitable observational choppiness. Unfortunately, though, it does not have much to say about the causally-efficacious 'hunk'[217] of matter and force that exists between these moments which embosses our three-dimensional canvas with physical properties—the very metaphysical currency required to establish change.

The relationship between discrete TPs and change is therefore not very clear, probably because of this independence from its physical composition. As Wasserman writes, "the mere fact that a banana has a green temporal part and a yellow temporal part does not suffice for a changing banana." (2006, 52) There is no necessary bond between time slices that provides for persistence, whilst there is no apparent relevance of these discrete TPs to the allegedly-persisting whole. And, this is the cause of our apprehension about discrete TPs, due to the insistence on understanding them as one would understand spatial parts. It is the dulling of what is an already dull, probably poorly-understood and possibly misapplied doctrine of general relativity.

Conceiving of time as an additional, fourth dimension is already an enormous challenge for the human mind. We can see that a Euclidian object has length, height,

---

[215] I would go as far as to say that what we call 'time' should be described as a 'mammalian-like perception of atomic and sub-atomic decay.' This is partially informed by current research on using time simply as a measure of change, and not as a separate dimension. One example is Sorli, A and D Fiscaletti (2012) "Special theory of relativity in a three-dimensional Euclidean space" Physics Essays, Vol. 25.

[216] This may be to say that time's ontological status is in question because of its inherent ontological dependence on whatever aggregate of matter and force we are observing.

[217] 'Hunks' of matter are a particular device used by Heller, as we will see in §2.3.

and width, and we can move along in that three-dimensional space as we move above, below and around the object. But to conceive of a Minkowskian dimension as being added on to the standard three requires us to *imagine* it so, so that we might mentally invent a possible observation that would demonstrate it. This mental imaging is necessary because of a) its non-linearity and b) its inseparability with matter and force.[218] The average human observer, with no technological assistance, typically perceives change of common objects as a directional and often linear process; and, whilst this (potentially teleological) view is dangerous, it is also embedded within our language on everything from bananas to galaxies—making it very difficult to formally conceptualise a general account of change. Language does not generally allow us to describe the incessant flux that flows in, through and around the very objects we are trying to use as hard metaphysical currency. Hence, drawing a fine distinction between time and that which it may very well be a property of is, so to speak, going in the wrong direction. Though the original intentions of the Einstein-Minkowski developments were to demonstrate both a) reference frame relativity and b) the juxtaposition of time onto matter and force, the outcome for some philosophy has been the over-emphasis on (and perhaps selectivity or biased preference for) the 'dimensionality' of time instead of its 'continuity'.

Most computer games use 3D space to design and arrange cells (like a room) and the objects inside them: moving this character over here, putting this wall there, turning this object slightly in that direction, etc. The objects are suspended silently in the 3D space waiting for the manipulators to find the right size, orientation and location. They are frozen in the cell, and the 'camera' can be moved near, far, and around these suspended, frozen, three-dimensional objects. The camera is moving, but not the objects. When the design is ready—after all the objects in a cell have been satisfactorily manipulated, and the file has been saved—the cell is added to the world-space. The game is then activated, and when we enter the new cell in-game, we are no longer using the camera to move. We (whatever character/subject) are actually moving, and the animated objects that we had designed and placed in the cell—which

[218] Also, Minkowskian time is understood as non-linear relative to Newtonian directionality and reversibility, which is a major focus of Prigogine's work (occasionally with Isabelle Stengers, e.g. 1984, but also 1997). The inseparability issue concerns the relationship between gravity and space-time, where the curvature of the latter is affected by the matter and force travelling through it.

had only been frozen in our 3D design mode, when we were giving behavioural instruction to these eventually-animated objects—are moving, too. We went from three-dimensional instantiations to four-dimensional dynamism, where two or more objects can interact, and structure and agency begin to blur. In such an example of a simulated, virtual reality—during either design or game mode—time is tractable, freeze-able, (one-) dimensional and discrete. There are behaviour protocols, cell boundaries, finite variables, and closed systems. There is no actual creativity, and there is no actual novelty.[219] Metaphysically speaking, there is no change, microphysical or autonomous, over time. Time is, in this case, discrete, and it is the facilitation of and the lubrication between the complex executions of binary protocols. It is the conveyor belt on which assemblages are created and destroyed. *Discrete temporal parts* would then make sense in this virtual world, as they would be useful for explanation as well as description of static, discrete wholes, agents and structures. Unfortunately though, discrete TPs seem to be limited to this kind of artificial and simulative application.

4.2.2. Sider & stage perdurance

Sider's stage perdurance, [220] which bears a resemblance to the time-slice approach, is based on the following definition of a TP:

> The temporal part of [an object] *x* at time *t* is sometimes defined as the part of *x* that exists only at *t* and has the same spatial location as *x*. (1997, 204)

This is similar to the discrete TPs from the last section, in that the TP is exclusively associated with one instantiation. Sider's definition also suggests that the TP at t_1

[219] Especially in open-world video/computer games—where a character can explore freely and is not bound by a narrow set of available actions—there is some question of emergent properties, such as when certain groupings of code produce strange effects. This kind of unexpected effect is another example of weak emergence: the designers simply overlooked the possibility of event X due to object Y interacting with condition Z. Though designers try very hard to find and fix these strange effects, or 'bugs', the complexity of the programme often means that they are inevitable.

[220] In trying to discredit endurantist three-dimensionalism (1997, 212), Sider casually suggests that, in the event that "scientists discovered that subatomic particles are constantly in flux," the endurantists might have to abandon their position. Written in 1997, there really is no excuse for completely missing the point: everything we know about the subatomic world—and have been discussing since Copenhagen—is in flux and fundamentally uncertain.

shares the same space that the object did at t_1, which is a small improvement from the general implications of the discrete TPs in that it recognises the need to connect—at least in some vague way—with the matter and force that compose the (spatial) object at the same instantiation. But our optimism fades as we see Sider's brand of 4D—not four-dimensionalism *per se*, even though his book is labelled and packaged as such (2001)—partially captured here:

> When you touch a person, you only directly touch a part of that person—the hand, say. According to the four-dimensionalist, there is another sense in which you only directly touch a part of the person. Even if you could somehow touch all of a person's spatial parts at once, you would still fail to touch all the person, for not all the person is then to be touched. To touch all of a person you must hold him in an interpenetrating total embrace from his birth until his death; only thus would you have access to all his past and future temporal parts. (2001, 2-3)

Obviously no one ever could or would want to hold another person from birth to death—excluding, perhaps, the parent of a sick infant. The fact that he chooses the action 'touching' to demonstrate the validity of TPs is an appeal to a singular common human 'sense' that curiously does not require a conscious mind.[221] The serious objections are quite obvious: there is no 'all the person', and there never could be.[222] Furthermore, he claims that in order to 'touch' this 'all', one would need an "interpenetrating total embrace." [223] Whilst the appeal to the irreducible is to be noted, this interpenetration is as it turns out precisely what he argues against in his *Thesis of Temporal Locality*:

> Necessarily, for any object x, and for any non-empty, non-overlapping sets of times t_1 and t_2 whose union is the time span of x, there are two objects x_1 and x_2, such that i) x_1 and x have the same parts at every time in t_1, ii) x_2 and x have the same parts at every time in t_2, and iii) the time span of x1 = t1, while the time span of $x_2 = t_2$. (1997, 204)

[221] A deep irony lies in the fact that no two hands ever really 'touch' as we like to imagine: the electrons in each hand repel each other such that there will usually be a distance of ~10^{-8} cm separating them.
[222] As we have been seeing throughout the work, there is no complete catalogue of properties, and, for this reason, there is no 'all'—and by extension, no 'one'.
[223] This embrace seems to be a misguided *reductio ad absurdum* in an attempt to shame the endurantists. He may be trying to goad them by using a Bergsonian or Whiteheadian term, but he has already stacked the deck as two TPs or objects may not overlap or share the same space. He is basically chiding 3D views for being that which his own philosophy prevents—or so he may think.

This 'thesis' is rather confusing, controversial, and quite possibly self-defeating. It suggests that *X* exists as x_1 at t_1 as well as x_2 at t_2. The duration (D) of *X* is the non-overlapping set of t_1 plus t_2. All in all, it would look something like this:

- ***(I)*** $X = [x_1]_{t1}$ — · — $X = [x_2]_{t2}$ $\rightarrow$ $X = x_1 = x_2$
- ***(II)*** $D_{t1} = D_{t2}$ — · — $D_{x1} = D_{x2}$ $\rightarrow$ $D_X = D_{x1} + D_{x2}$

The consequences are intriguing. We are not informed on how *X* comes to be, how *X* comes to be known, or on what *X* is constituted by other than its arbitrary parts x_1 and x_2. Yet *X* is both x_1 and x_2—assuming x_1 exists at t_1, and x_2 exists at t_2—the implications of **(I)** being:

- *(i) X is either of its parts.*
- *(ii) X is both of its parts.*
- *(iii) If X is either, X does not persist.*
- *(iv) If X is both, X endures.*

Are we inferring from the 4D author's thesis that his 'perdurance' is actually presentist [as in (iii)]—or worse, 3D [as in (iv)]? He may rebut with the claim that the existence of TPs (though not yet observed or demonstrated) prevents us from saying that *X* wholly persists, which is necessary for endurance. But, then we would be left with discrete TPs, and that is not favourable. Intriguingly, discrete TPs are also instantiations—or at least they are required to be, based on how we have described them in §2.1. Yet, Sider's TP's have durations, and they are, amazingly, equal to that of their corresponding ('matter-and-force' or 'spatial') parts mutually instantiated. *X*'s duration is then the sum of the durations of its two spatial parts. As soon as he claims that TPs have durations, many other claims are jeopardised. We know that *X* consists exclusively of two parts (x_1 and x_2), and that *X*'s parts are of equal duration, where *X*'s duration is from t_1 to t_2. *X* must therefore exist both at and between t_1 and t_2. If t_1 and t_2 were to have durations, they could not be instantiated—meaning, our observations would have been nullified—and Sider is creeping even closer to 3D. For t_1 and t_2 to have durations, x_1 and x_2 must also have durations. If they do, then they persist, and we are essentially dealing with proper discrete parts of an enduring whole.

Ultimately, Sider's 4D requires discrete, non-overlapping temporal parts in the context of discrete objects. If we incorporate the feature of his TPs where they have durations, his entire attempt at applying Minkowski to ontology and persistence is in question. If we do not, we are perhaps looking at a position that seems to be struggling to detach itself from a targeted endurantist audience—that still finds itself tempted by closed systems, universals, and even observation-independence. Heller, the other notable 4D theorist, similarly uses quasi-enduring objects and discrete TPs, but offers a unique way of envisaging general persistence.

4.2.3. 4D worm perdurance

In *The Ontology of Physical Objects* (1990), the author immediately discards the endurantist picture as well as the vision of an enduring spatial 'hunk' of matter. He sees matter not as filling up regions of space, but rather, regions of space-time. As such, "a physical object is the material content of a region of space-time." (4) [224] He then repeats the 4D mantra of seeing temporal parts as one sees spatial parts—which necessarily and problematically leans toward discrete-ness. Referring to persisting, 4D hunks of matter as 'worms' is partially based on the time-slice view, as shown in Figure 4.2.1. The object exists at multiple instantiations, and whilst the matter and force present at one may not be at another, the object persists due to the contiguity of these slices. It is the presence of temporal parts that makes the object persist. Because its matter and force changes over time, the object does not endure. But, since the object at t_1 is said to be the same as the object at t_2 in virtue of this contiguity, it can be seen as the same segmented worm: replace the oblong object in Figure 4.2.1 with that of a common earthworm, and the label makes more sense.

In his attempt to show its weaknesses, this is how he sets up the 3D position:

> To see the contrast clearly [between a spatially and temporally-extended object], consider an object that is created at noon and destroyed at one. If we think of the object as three-dimensional and enduring through time, it would be appropriate to say that the object

[224] Now, the phrase 'region of space-time' must be understood as exceedingly vague, and ultimately just an instrumental expression for an approximated, though ultimately open and dynamic three-dimensional area being monitored over some duration.

> exists at different times; the same objects at noon and at one. Such an object has boundaries along only three dimensions. The whole object is that hunk of matter that entirely fills up those boundaries. The whole object, therefore, exists at noon and still exists at one. (1990, 4-5)

This exposes some of the flaws, but he cannot seem to break away from the discrete problem. Of course, we have no reason to consider any object as being created *ex nihilo* in a metaphysically-significant sense, but rather re-aggregated through physical means with only a vague duration and an inconsistent material composition. Even if an object had been created—or even if we change 'created' to 'spontaneously aggregated', the instantiation of 12:00:00 *ad infinitum* (henceforth using the symbol °) is meaningless and non-falsifiable, to the point where we cannot say exactly when it came into existence without capping the number of decimal places (and there is the same problem at 13:00°). Only the person who accepts that the object was actually created at 12:00° and destroyed at 13:00° could continue, because the statement presupposes that there is such a thing as 12:00° and 13:00°. These are instantiations, and not real things or real occasions.

There are several other points on this passage. (1) If the object is physical, it is already spatial as well as temporal, because, as we have been discussing, time is better understood as property of matter and force (which implies 'in space'). (2) A refined physicalist and empiricist (and therefore nominalist) approach cannot grant an essential unity at 13:00°, because that would require a complete catalogue of properties at 13:00°. (3) Any physical object at t_1 is 3D, as it has been instantiated—as it is the same in the design mode from §2.1. (4) In order for the 3D object to endure (to wholly persist), it needs to be unified and of uniform composition at both 12:00° and 13:00°. Since this is impossible to know, let alone falsify, Heller should not assimilate three-dimensionality—the only canvas we have for observation—with endurance. This is common of most of the 4D crowd: they use instantiations, and then suggest that surviving objects must be wholly present. In a nutshell, they have a built a system of 4D upon a foundation constituted by the weakest elements of their opponents, the endurantists.

Overlooking for a moment the ontological nature of his claims for discrete temporal parts, Heller spends a large section of the book (1990, ~32-47) using the example of

Manhattan as a spatiotemporal hunk of matter and as what he calls a "conventional object"—a loose, nominalistic label for an approximated region of space-time.[225] The section contains questions of historical identity and persistence, changing properties, and spatial boundaries, where all the key metaphysical questions seem to entangle themselves. How do we know what 'Manhattan' is? Is it not the same as well as different from what it was 200 years ago? How can we reconcile coinciding spatial and temporal objects? Is accepting temporal parts the only way to have an idea of what it is? These are all good questions, and all relevant—but he also litters the debate with talk of essential properties, *ex-nihilo* objects, and why conventional objects are different from generic objects or statues-from-clay objects. Like most of the 4D authors, it seems as though they are fighting a battle with archaic notions that could easily just be ignored; and yet, in order to contest some of these notions, their own claims are muddled by some of the same errors they are trying to overcome.

4.2.4. Refining worm *persistence*

It seems that the 'hunks of matter' metaphor could be improved. Instead of persisting unified hunks, with *hunk-at-*t_1 discrete time slices with ontological priority, why not consider rough hunks of matter with volatile, spatiotemporal parts? 'Manhattan' is a label for a hunk of land with a long history, with no precise boundaries or intrinsic or essential properties. People and things come and go, interact with other people and things, and necessarily and inevitably change the physical and social features of whatever region of space-time we want to roughly designate as 'Manhattan'.

If we shrunk the borough to fit in the palm of a human hand,[226] or if we simply wanted to pretend that any clump of dirt was a bustling bacterial city, imagine that I am holding it for five minutes. Just in that short amount of time, the properties of the hunk have changed. Some miniscule particles have been carried away by the wind. The moisture

[225] The term 'conventional object' curiously includes most proper nouns (excluding humans, who apparently are too special to be lumped in with other named things!) but not common nouns, like trees and tables. The obvious question is, if 'Manhattan' is a conventional object, what is 'conventional object', if not a conventional object?

[226] Let us, for discussion's sake, ignore the issues of a) how much of the earth's crust to take with us into the hand, and b) the topographical precision of its peripheral boundaries.

in the air has made some pieces stick together. More energised particles have caused some of the moisture, which had been keeping other smaller clumps (within the greater hunk) together, to evaporate—making them vulnerable to crumbling. Tiny cardiorespiratory pulses, partially affected by the body's gravitational pull on the ground surface, prevent me from keeping my hand perfectly still, affecting the integrity of the greater hunk as well as smaller, dense sub-hunks. Local temperature and humidity affect the glandular secretion of oils and chemicals on my hand, causing the amount of friction between the clump and my skin to fluctuate—which, in turn, affects the consistency of the hunk. The vibrations on the surface of the ground upon which I am standing—which even causes my hand to shake slightly—also cause the weaker molecular bonds to break, making the clump of dirt less consolidated than when I first picked it up.

This is only a brief list of some of the microphysical changes,[227] and there is, ultimately, no limit to the number of changes relevant to how we construct the object's quiddity. The hunk of matter and force that we hold in our hand is precisely *not* the same hunk from one moment to the next. It is the fact that we are calling it a hunk that hopefully makes it clearer: a hunk has no name, it has no clear shape, it is composed of loose bits, with no purpose or design, arbitrarily, hastily, and probably clumsily extracted from a random patch of earth. The reader is invited to recall the passage from Chapter I, §4.0, which describes a "random grab" of objects: the thrusting of one's hand into a collection of unknown things and blindly removing what he hopes to be (or wants to see as) a singular unit of analysis. The hunk of dirt is loose and fragmented, with small, medium and large-sized particles falling to the ground as one gradually stands back up with the hunk in hand. Its properties and constituents never stopped changing, such that we should even be hesitant to consider it 'change' at all. The difference, incompatibility, and sameness conditions from §1.2 above are all, to say the least, blurred.

One objection to follow would presumably be that the 'hunk' example is an easy target for showing disunity and dynamism, and that hunks of dirt could hardly be considered

[227] And, more importantly, there could never be a complete list of microphysical changes. But, that does not mean that all attempts at explanation are futile. It just means that explanation is never complete, because the description never is.

as 'units of analysis', let alone as 'objects' or 'wholes'. Numerous examples, though, throughout this present work, should have shown the material inconsistency and environmental porosity that many, if not all, of our most common objects have. Suppose that we take that clump of dirt and put into an air-tight container: is it suddenly an object, a unit of analysis, and/or a whole? Why are we more likely to believe that by 'sealing a container', we create a closed, unified system? Is it because we can give it to another person as a gift, or perhaps because we can place it neatly on a shelf? To claim that the hunk inside the container undergoes no physical change would reject all physical evidence and explanation that has been laid before us. So why does its encapsulation suddenly convert it into a whole from a heap?

Calling something a 'hunk of matter (and force)' is profoundly important. We know that 'hunks' lack unity and material consistency, and yet, they are weakly individuated. One hunk is one loose object. Whilst the container (from the previous paragraph) is neither unified nor maintaining material consistency (not to mention local conditions, which would feed back into its consistency), it, like the clump of dirt inside of it, can be a unit of analysis. It does not matter if we take the clump, or the container with the clump inside. Both of these constructs are potential units of analysis. We have already decided that they cannot be wholes, let alone closed systems. But, they can be studied. Let us say that we want to study the dirt, and not the container, because we made the container, and we already know many of its properties (at least at t_0). We can study the hunk, and all the properties that we ascribe to it—over time, and at t_1—relate to that which we are studying.

The investigator, knowing that she is studying a loose clump of dirt, will not attempt to ascribe essential properties to it. There can be nothing that binds the particles, because the particles are not bound together in a metaphysically-significant sense. Because of this, the clump can have no autonomy, and no autonomous causal powers. She will try to ascribe generic properties to 'the sample', by conducting, for example, elemental testing which shows that the sample contains traces of phosphorus, nitrogen, and potassium. There also may happen to be traces of nickel, for example; and that would make sense, because the sample was taken from an area known to have nickel deposits. The rest of the sample, at t_1, contained the usual range of gases, liquids, solids, and organic molecules, as well as bits of 'rock', 'dust', animal hair, a few small insects, and a huge collection of bacteria. Very little had changed at t_2, although the moisture in the soil had evaporated, and much of the bacterial life had died. The point is: we can study the 'hunk', just as we can study an organism. When we know that we are studying a hunk of something, we just collect all the properties that we can about the hunk—the sample. That collection of properties is our non-exhaustive, fallible, mutable, parsimonious, pragmatic, instrumental best guess as to what the 'hunk' is. This is all that we could ever hope that a quiddity could be.

4.3. Persisting quiddity

A spatiotemporal hunk of matter and force persists, whether or not we want it to, or whether or not we want to believe that it does, or whether or not we observe it persisting. This is one thing that we can say with 'hunks'. It is simply a cluster of particles with non-uniform composition: those particles existed before we had grabbed it from the ground, and they will exist after the hunk has been broken apart or disintegrates. The things that composed the hunk were already persisting, and therefore, to say that the hunk persists is to state the obvious.

4.3.1. Properties into quiddities

To demonstrate this, drop down a level of analysis, and exchange the clump of dirt in the hand with a clump of particles. Some bits stay longer than others, some flitter off, some fuse, some disaggregate, and all of this happens with local conditions fluctuating in, through and around them. At the level of the atom or even the molecule, the spatiotemporal 'hunk' is a just a node in a whirlwind of microphysical dynamism—to the point where the hunk can be eroded down to when it is indistinguishable from the whirlwind around it.

Imagine that I place a handful of sand on a gravel sidewalk at t_1. A gust of wind comes through, and the shape of the heap of sand, which had only been approximated, begins to change noticeably. The particles are being moved around, and as the gusts continue, the heap is being gradually reduced to a barely-identifiable thinly-aggregated pile. Eventually, the wind succeeds in carrying seemingly all of the particles away, and 'the heap' no longer exists at t_2. Because the particles were light enough, and not bonded sufficiently to the other particles, the wind was able to overcome individual grains' gravitational pull on the pile and the earth, and separate and distribute the grains according to its strength and resistance at various moments and places.

The handful of sand at t_1, as intuition tells us, may have some standard properties: perhaps, a diameter, volume, weight, mass, some colours, friction, etc.[228] We could, if we were feeling generous, ascribe these properties to the handful qua handful. We could consider the handful at t_1 as a unit of analysis H. However, we cannot consider it as such at t_2, because there *is* no handful at t_2. All of the properties of H, at t_1, are no longer ascribable at t_2. Whilst it does not make sense to say that there are no properties of H at t_2 because there is no H at t_2, it *does* makes sense to say that there is no H at t_2 because there are no properties of H at t_2. H is instantiated at t_1 because there are enough properties in a loosely-aggregated (but aggregated) hunk of matter and force at t_1 such that we can associate these properties together in a condensed format. The particles that H had been composed of at t_1 still exist, but they are not associated in the same way. They have been separated, and we are unlikely to see evidence for any unity of, or a bond or association between them. The unit of analysis simply ceases to exist at t_2. So, is localised property ascription the prerequisite for object individuation in a refined physicalism?

The sand grains that composed H are still around, and the loose shape and other scattered properties that we had attributed to H had never been metaphysically significant. How can the unit of analysis disappear, if everything composing it is still there? The answer is properties. Properties were what enabled us to call a vague heap of sand a unit of analysis, properties allowed us to ascribe other properties, properties allowed us to determine that sand can be blown away by a strong wind. Our quiddity, that which H is, or H's what-ness, does not need essential properties, essential or proper parts, or universals. Our quiddity is constructed from properties.

[228] All of these are, as we have been saying throughout this work, approximated and observation- and condition-dependent.

4.3.2. Persisting properties into persisting quiddities

Let us say that the winds in our example stretch out from t_2 to t_4, where we place the handful of sand (H) on the ground (with no wind) at t_1, and the wind is not strong enough to visibly affect the composition or arrangement of H until only slight changes are visible at t_2. H, at t_3, has lost roughly half of its composition, and is entirely blown away by an increasingly-strong wind by t_4. If we know that the some of the particles that existed at t_1 have been blown away by t_2, why are we still referring to the altered pile of sand as H at t_2 and after? It is not the same H, but we are still calling it H. There must be something about *H-at-t_2* and *H-at-t_3* that reminds us of *H-at-t_1*. Most perdurantists, as mentioned above, would step in and say that there can only be *H-at-time-T's*, and no 'real' H. But since, as we have said, there cannot be any ontological priority given to instantiated time-slices, and since 3D matter/force aggregates can never be separated from their temporal embossing of our 3D 'canvas', *H-at-time-T* is just for our convenience. We created the idea of a frozen moment, and as such, it is for us to determine how we derived H from an incompatible or at least inconsistent set of {*H-at-t_1*, *H-at-t_2*, and *H-at-t_3*}.

Intriguingly, we were only able to recognise that there *was* a difference between *H-at-t_1* and *H-at-t_2* or *H-at-t_3* because of properties. The change is via, and is itself, properties. They are the only currency able to describe and explain these objects, events and other properties. The only reason we are even considering *H-at-t_1*, *H-at-t_2*, and *H-at-t_3* as evidence of persistence is because there are some properties found in all of these snapshots. They might include a) the colour of most of the grains of sand, or b) the pile's seemingly general topographical consistency. In a), there is the very flimsy and superficial ascription of a uniform colour to whatever grains exist or remain at t_1, t_2, and t_3. I have discussed the problem of colour with regards to optical and molecular physics and local lighting conditions, so we need to tread carefully. There is, however, the idea that an observer may ascribe a self-described limited range of colours, or a (human-designed) supercomputer can 'confirm' that there was an insufficient degree of change in the grains' chemistry as well as an insufficient change in local lighting conditions, for example—such that we may lean towards saying that certain grains have lingered between observations. In b), we might notice that the sand had been laid at t_1 upon a certain large block of rocky concrete with visibly-discernible borders, and that *H-at-*t_2 and *H-at-*t_3 (where neither are composed of particles that have already flown away) are both still confined to that same block (or roughly the same area as it).

In both a) and b), we are dealing with properties, but they are properties that are heavily-prone to human error and local interference. Ascribing a 'general yellowness' to a group of particles, especially when some of the particles are not visible (because they are lying underneath other particles, obstructed from view), is of limited descriptive and explanatory power. The topographical contention is mildly useful, but it does not lend itself to precision, as it were, because the surfaces upon which the sand rests are subject to movement and compositional change. If we wanted to specify a location where the sand rests, in order to justify a claim of *the same H*, we could never be specific enough to be metaphysically significant. The sand moves, the gravel slab moves, the crust moves, the earth moves, the solar system moves, etc., and they all move and change in different ways with different micro- and macro-physical effects.

According to this stringent or perhaps overly-generalised standard, it may seem like nothing could ever meet this criteria for persistence. But, it is crucial to show how flimsy some of our standard property ascriptions are.

- ❖ *The house was red yesterday.* → *The house is red today.* → *It is the same house.*

The conclusion, that it is the same house, is meaningless, as there are problems of identity and persistence embedded in this example. We know that the house is at least micro-physically different—which means that the house is *de facto* a different house, because the house is exclusively physical. But, we do not even know what 'house' is (e.g. precisely what needs to be present at both t_1 and t_2 qua house), nor do we know what 'red' is (including what precise conditions are required for it to appear 'red'). We have some rough guess of what these things mean, and based on those rough guesses, we might be confident that, for example, the 'same' engraved table and chairs will be inside on both days, or that the registered house number is identical. And, that is all we have: a loose collection of properties, some of which linger through multiple observations. Those that linger are most crucial to informing that quiddity, but by no means are essential to it—and this is because there is no universal or observation-independent quiddity of any given unit of analysis.

4.3.3. Entia successiva versus entia per alio

Let us revisit Chisholm's *Person and Object*. He has three kinds of entities, the least interesting being the *ens per se* (let us call it 'Option 0')—where an object exists in and of itself, and whose members include human persons and a god of some sort.[229] The *ens successivum* (let us call it 'Option 1A') is that which is "made up of different things at different times," and is a "logical construction upon the various things that may be said to do duty for it." (98) With the Scholastic formulations mentioned in §1.4, the *ens successivum* is also that which has some, but not all, non-overlapping, discontinuous, and disconnected parts. There is, obviously, quite a large number of objects that would meet those criteria; however, he adds another kind of entity meant to be compatible with Option 1A, but not with Option 0, called the *ens per alio*, or, for our purposes, 'Option 1B'. It is outlined in these two passages:

I. […] Appearances would be paradigm cases of […] 'entia per alio' and what we might call 'ontological parasites'. They are not entities in their own right; they are 'parasitical upon' other things. And what they are parasitical *upon* are persons or selves. (51-2)

II. [Consider] the transfer of a shadow ('the shadow of his hand moved from the wall to the table and became larger but more faint in the process'). But a shadow is an *ens per alio*; it borrows its properties from other things (most notably from shadowed objects). (107)

We are meant to understand that an *ens per alio* is an object that derives all of its properties from other objects. Unfortunately, the best example that the author can come up with is that of a shadow. Shadows have long been used in thought experiments to demonstrate puzzles in almost all sub-domains of metaphysics, though they are hardly mysterious for optical physics. A shadow is the visible effect of an obstructed light source, a perceived region of space-time containing variations in lighting conditions. The word 'shadow' is a great example of an abstract object (e.g. 'happiness') commonly and erroneously understood and/or used as a concrete object

[229] Since we have no reason to think that we will ever observe an ontologically independent entity, Chisholm's *ens per se* is of no interest.

(e.g. 'chair')—and is a nice demonstration of the futility of many classic puzzles due to the imprecision of language and the obstinate yet implicit insistence on universals. Chisholm's decision to use a shadow as an example of an object that derives all of its properties from other things is perhaps a regrettable one, because it makes his argument for *entia per alio* much harder than it needs to be.

He makes the crucial mistake of saying that *entia per alio* derive their properties from *entia per se*—meaning people or god. First of all, all objects—including all three of his kinds—have been individuated via properties. So, whatever objects he places into the *entia per alio* category must also be individuated via properties. But, if properties derive from people or a god, there would be no observational interference or irreproducibility of investigatory results, for we could just agree on all properties, and could therefore claim to understand the entire universe.[230] Observations would always correspond to our descriptions and explanations, and actual natural investigation would, at some point, run its course. It is precisely because properties occasionally *do not* correspond to our observations, descriptions and explanations that properties clearly do not derive from *entia per se*.

Properties cannot be said to *derive* from the conscious observer, nor can they be said to have been generated *ex nihilo*. At the same time, though, they do not represent objective knowledge, as we have been saying, which means that there are no intrinsic or essential properties—at least in a metaphysically-significant sense. Colour, shape, length, speed, and even mass[231] are all subject to at least conditions, and in most cases, convention as well. Properties are the coded sensory data shared by conscious observers, because they are the yellow adhesive labels that we put on things to remind us of which selection of coded sensory data describes them. It is true that properties need language, or qualitative identity, and measurement, or quantitative identity. But, they also correspond to observation, and the sensory data that is processed during this observation-event is eventually translated into properties—whether by an

[230] This is, of course, precisely what most (codified, dogmatic) religion does or attempts to do. Leibniz, for example, and his quasi-Christian or deist sentiments, led him down this path.
[231] The reader is referred back to Chapter Three, §3.1.1 for more on this. Also, the property of 'number' is not essential either, because it presupposes unity and qualitative identity (see Chapter Four, §1.1).

individual brain or processor, or a community of conscious observers and/or their processors—that we can use and understand.[232]

We put the label ‘red’ on a ‘red apple’ to remind ourselves that this is what we agreed ‘red’ is. We did not send the reflected red light to our eyes and brain. The light *just came*, and our eyes and brains are fortunate enough to be able to process the information. We came up with a word to correspond to a group of things that all roughly reflect a similar kind of light under similar conditions. Whilst there is no definition or universality of red—because our eyes perceive only a small amount of the electromagnetic spectrum, and they and the objects and spaces they perceive will always be subject to local physical conditions—there can also be some agreement that, under normal and fairly constant Earthbound conditions, we can determine a range of what might be considered loosely as ‘red’ and ‘not-red’. Certain objects, largely because of their chemistry, will reflect red light under certain conditions, and typically this happens without us noticing. We would not say that the ‘red’ flower that grows in the forest is not red because nobody saw it. The chemistry is such that it reflects light that we recognise as ‘red’ under common Earthbound conditions—and so its redness does not require our approval to be that which we have been calling ‘red’.

Parasites are microorganisms that benefit from living off of another organism, though, in the process, must prove to be detrimental to the host organism. Throughout the book, Chisholm emphasises the person as an *ens per se*, despite the fact that, ironically, adult humans are so over-infested with microorganisms that it is even debatable whether they are living off of us or vice-versa.[233] In other words, bacteria living inside our bodies are so important to many essential general anatomical and biochemical functions that it is not clear which party is the parasite—if either. From either position, neither is an entity in and of themselves, and the bacteria and the

[232] Properties are the tumultuous marriage of Kantian *phenomena* and *noumena*, in a sense. There is an inevitable incongruity between all physical phenomena, our sensory data, and our coded sensory data. Cybernetics and artificial intelligence, for example, have always aspired to resolve the conflict between the latter two, but there is no hope for the reconciliation of all three—most importantly because the cosmos is an unknowable thing with unknowable and indeterminate causal relations.
[233] There is an interesting and helpful passage on this issue in Dupré (2008, 15).

human 'organism' are mutually dependent.[234] Parasites generally carry a negative connotation, though they really represent a fundamental feature of at least the 'biosphere': nothing is closed and an *ens per se*, nothing is independent, and everything, even if only to some minute degree, is dependent on everything else. Ontological parasites, qua *entia per alio*, *share* properties with other objects—not necessarily in a symbiotic sense, but in entangled, disordered, interactive, dynamic and diachronic ways.

So, might we see a parallel between Chisholm's property derivation and Searle's property-ontological reduction (1992, 70)? This would mean that, for every object whose properties are derived from other objects, we are to reduce these parasitic objects to the properties of their parts. Searle's description of this kind of reduction, however, does not necessarily require a dynamic, interactive composition, for he is happy to reduce 'heat' to the mean kinetic energy measurement of molecular movements—and leave it there. A bottomless *ens per alio*, lacking ontological primitives, cannot stop there, because molecules have properties, too, and we can open molecules up and see what is inside. It is conceivable that the *ens per alio* is meant to concern non-spatiotemporal objects, as Chisholm tries in this section to avoid declaring. In which case, all ideas (for example) will evidently derive all their properties from other objects. But, despite this obvious statement, ideas are not included in a physicalist ontology, nor do they have any spatiotemporal properties. Hence, either Chisholm is going in the wrong direction, or has misunderstood the importance of the *ens per alio*. I believe it is the latter, because property dependence, as we have been alluding to throughout this work, is interchangeable with ontological dependence. This is due to the fact that ontological dependence can only be claimed based on property dependence (or property explanation), since properties are the primary currency in a refined physicalism and empiricism—and the sole qualifier for the other currencies of events, objects, and other, more basic properties.[235] If ontological dependence was

[234] This example of bacteria also weakens the notion of a unified person—though this is outside the scope of this work, which, again, has deliberately avoided issues of persistence of people, personhood, cognition, and volition. The dependence on the bacteria is only one demonstration of a human's ontological dependence: other crucial dependencies include that of constant transmission and exchange of heat, light, oxygen, and nutrients, all themselves dependent on suitable local conditions.

[235] Basic properties are the initial properties yielded from preliminary, unaided, unexplained, primitively-coded, observation. We look at a red apple and a red flower, and we ascribe the property 'red' to the both. Later we discover that 'red' as a property is not very clear, and dependent on several factors. Objects, events, and basic properties are qualified by other properties, which are generated by

disassociated from property dependence, many objects could be considered as ontologically *independent*—and we know this is an unfavourable position to find one's self in.

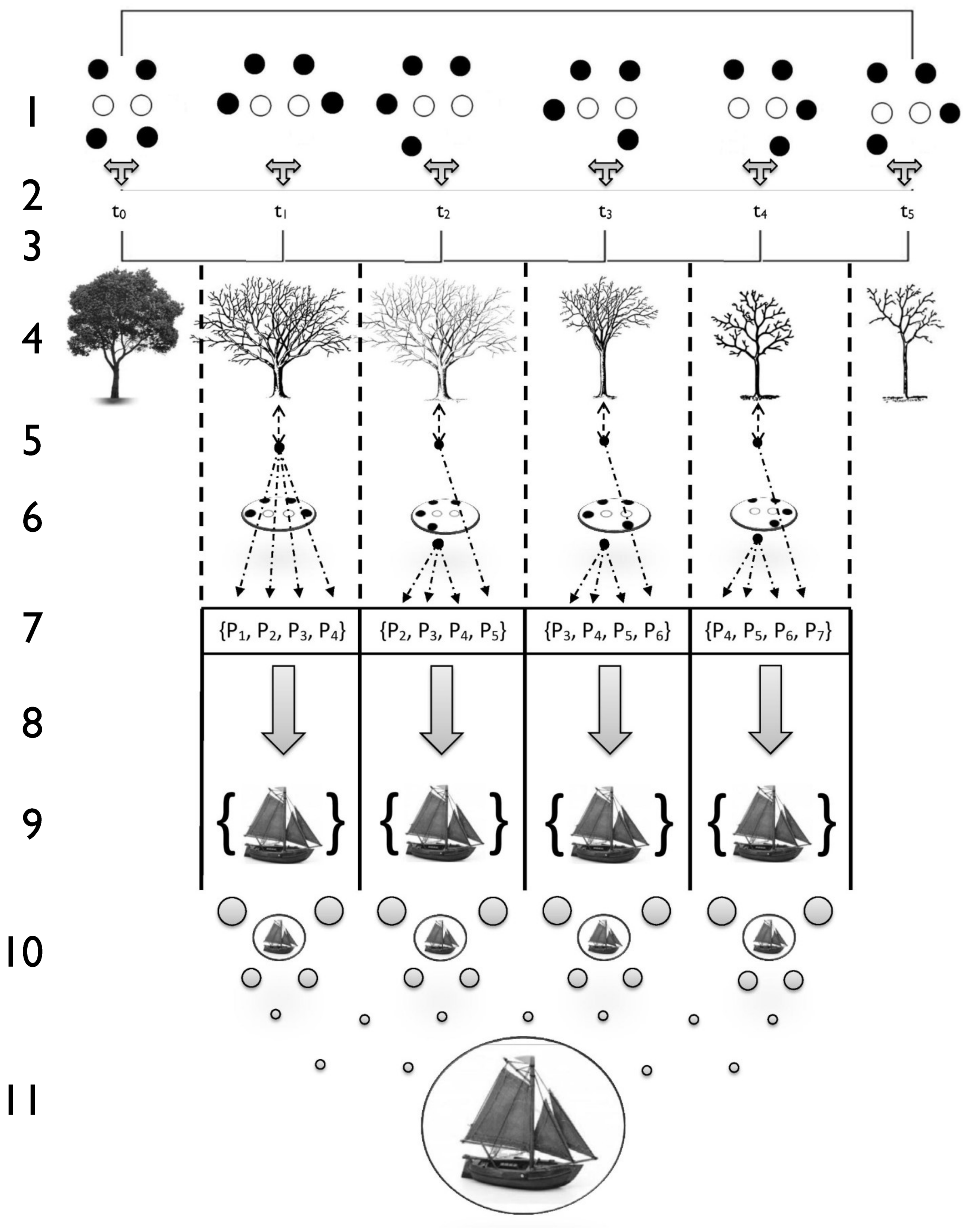
When we examined the *entia successiva*, we found that many, if not all units of analysis, were constituted by some (but not necessarily all) non-overlapping, discontinuous, and disconnected parts, and were, in Chisholm's words, logical constructions upon the various things that may be said to do duty for it. In other words, *entia successiva* have dynamically-persisting compositions with some lingering properties, and there could not be, evidently, anything recognised by physicalistic-empiricist investigation that did not meet those criteria. So, what could we hope to do with the *ens per alio*, in light of the all-encompassing *ens successivum*? The *ens per alio* points out the vagueness of properties *per se*—which would be instantiation-based—that would give us reason to challenge the ontological status of wholes (and their parts), objects, and units of analysis *at a given instantiation*. The *ens successiva* points out the vagueness of persistence via properties—which would be over multiple instantiations—that would give us reason to challenge the ontological status of wholes (and their parts), objects, and units of analysis *over the course of several observations*. It seems that they could be understood together as laying a solid foundation for a refined physicalist ontology and eventually a new naturalistic metaphysics. For now though, let us see how *The Ship of Theseus* fares in a reformulation of the classic problem with these two kinds of objects in mind.

4.3.4. *The Ship of Theseus* qua ens per alio qua ens successivum

We know now that we never could have a complete catalogue of properties for the ship (S) at any of the instantiations, and this prevents us from claiming endurance. The question is whether and how S persists, but—with a slight contortion of Chisholm's taxonomy—we may be able to see how S can do so within a refined physicalist framework. Figure 4.3.1 below is meant to illustrate the various processes necessary

observation and `eventual description—which, when explained, yield better properties with better observations, better descriptions and, eventually, better explanations.

Figure 4.3.4.1: The Physicalistic Persistence of The Ship of Theseus



for the persistence of a whole with parts, leading to the construction of an *ens per alio—ens successivum* hybrid. The reader will see eleven numbered lines in the diagram, with each corresponding to specific events, as outlined below.

4.3.4.1. Diagram Explanation

(1) Our physicalist base recognises only matter and force as ontological quasi-primitives, but not in one discrete or atomised form. This means that dynamic composition is best identified and explained from the bottom-up through time. This also means that everything is aggregated based on condition-dependence, and that properties of these aggregates can and must be reduced to or explained by other properties (including those of objects, events and conditions, and their properties). The diagram shows slightly different arrangements of the matter and force, to emphasise the irreproducibility of any precise configuration (which could not be known precisely).

(2) The triple arrows indicate the free, non-linear flow of these particles that will eventually find themselves contributing to the composition of our unit of analysis. (3) I am using this linear model of time to reflect the instantiations that correspond to observations. The matter and force that composes the units are not restricted by these instantiations, which are only being used to show different observations with different properties. Matter and force existed before we began thinking about or building the ship, they exist and have properties in between these instantiations, and they will continue after the observations of S cease at t_4.

(4) The particles find themselves composing, in part, the trees that will be used to construct the proper parts, beginning at t_1. The tree existed before this, both as tree and a seedling, etc., and will exist after t_4, but the composition and properties will necessarily be observationally different. (5) Here, there is an emphasis on the transmission of particles almost exclusively from the tree at that instantiation, and not from another tree or another macro-level object. The line is dotted, however, indicating that the process of this transmission does not happen in a vacuum: microphysical changes are inevitable before and during the cutting and sawing. For example, the

entire tree's chemical composition is radically altered the moment the outer bark has been breached: the system begins to work to maintain a balanced distribution of nutrients despite this puncturing.

(6) The wood from the tree, already altered (or still compositionally dynamic) before being cut, is exposed to more, different particles and micro-level changes in local conditions as it is being processed. The planks needed for S need to be smooth-surfaced and sharp-edged, which results in massive compositional changes and, for example, evaporation of water and loss of minerals. The wood, in a sense, becomes increasingly weaker the more the fibres are exposed to the open air. At t_1, a new set of planks is needed, and so all of the materials that will constitute the proper parts of what will become S are taken at the same time and exposed to a relatively narrow range of conditions. Subsequent instantiations involve going back to a source of wood and repeating the process for a replacement plank. The source of wood, whether or not it is nominally the 'same' tree, is a different source, as it has inevitably undergone vast microphysical changes. The new plank joins the old planks, but even the old planks are qualitatively different than they were at the previous instantiation(s)—as they also have been exposed to fluctuations in local conditions as well as molecular and chemical changes.

(7) Up until 'Line 7' in the diagram, only dotted lines have been used—with the exception of the number line, which is instrumentally linear to mimic the (imperfect though instrumental) linearity of our observations. The dotted lines represent the porosity—and, by extension, the uncertainty—inherent in any unit of analysis, as well as any object, property, and event ascription. There is only the illusion of the closed and the static, which is coordinated with our observations, and necessary for our descriptions and explanations. Once we arrive at this seventh line, we see the beginning of the individuation process. The wood from the tree is cut, trimmed, shaped, smoothed, and treated, and turned into what appearances may present as interchangeable and identical proper parts. In order to consider S at t_1 as individuated, planks one through four also must be considered as individuated. The set $\{P_1, P_2, P_3, P_4\}$ must be discrete and bound together as the exclusive composition of S at t_1.

(8) The downward arrows show that, when we have exclusively discrete objects, we immediately begin to *mereologise* S. If S is made of proper, discrete parts, and is interchangeable with the set $\{P_1, P_2, P_3, P_4\}$, then S is a whole, from which nothing or no part is missing. The project vision, so to speak, was needed to determine why and how the planks were formed and put together, so the idea of the whole came before the fabrication of the parts. But in order to manifest the whole—to bring the spatiotemporal whole into existence—proper parts were needed. In 'Line 7', we see the aftermath of a decision about what materials and elements were necessary to build a ship. The components were assembled, set aside, and laid on the ground. They were everything that was needed, and as such, are the set $\{P_1, P_2, P_3, P_4\}$. (9) The set becomes the whole at each instantiation, so we have four sets and four wholes.

(10) Now the vertical lines fall away, and our physicalist-empiricist mereologising is interfered with by memories and previous notions of 'ship'. We have four wholes that are slightly different, but bear some general resemblance to the original plan as well as to previous notions of 'ship'. (11) We are likely, whether at some point after t_5 or before, to declare that there is an insufficient amount of change in S to deny persistence *of some kind*. There *are* overlapping proper parts, such that our catalogue of observations indicates an adequate kind of continuity, and we can 'feel' confident that S persists qua whole *simpliciter*.

4.3.4.2. Diagram Analysis

The Ship of Theseus was always meant to be constituted by proper parts, so challenging every aspect of S's persistence on the grounds of microphysical variation or metaphysical vagueness will typically be met with rolling eyes. The important feature of this diagram, I think, is that the even when we grant proper parts at one instantiation—which is already a generous offering, considering the high stakes of persistence—we still have to justify those proper parts' persistence to later observations. According to the original formulation, S cannot persist without its proper parts, and there are serious problems with the account of persistence of S via proper parts. We need *that-which-we-have-been-calling-S* to persist *in some way*, because we are using this classic puzzle to demonstrate the problems of any persisting unit of analysis—which, again, is necessary for description and explanation.

Intuition might suggest that the objects which persist must also be at least integral wholes, largely because they have at least some overlapping *parts*, as in the *ens successivum*. Any temporally-overlapping part must also be a discrete spatiotemporal part, since it would need to be individuated in order for us to determine that it is overlapping (present at multiple instantiations). What is of equal, if not greater importance, however, is the overlapping *properties* of those parts, that not only must successively re-appear, but must also be manifested by suitable conditions—as in the *ens per alio*. Parts, especially proper ones, have properties, and are composed of other parts, which have properties; but, properties alone do not make a whole. The properties have to be bound to—or applied to—objects, events, other properties, and conditions, and their properties. For the purposes of description and explanation, we are going to need a quiddity, and a quiddity that survives multiple observations—which is to say, no discernible properties are omitted.

The Ship of Theseus must be seen as being both an *ens successivum* and an *ens per alio*, for its persistence, anchored by a quiddity, depends on it being both. The only way to generate a quiddity—the what-ness of S, and that which permits the individuation of S—is to reconcile the property-dependent part with the part-dependent property. From this, a whole can change and stay the same. So, how can this reconciliation be achieved? The answer, finally, is a reconceptualisation of emergence.

4.4. Manifesting dynamic and diachronic wholes

Throughout this present work, I have emphasised the importance of seeing wholes not as wholes *per se*, but as wholes qua units of analysis. This should, as I have tried to demonstrate, eliminate any confusion about essential parts, essential or intrinsic properties, or essential, objective quiddities. It also reinforces the neo-physicalistic and, by extension, nominalistic terrain upon which we are analysing the wholes qua units of analysis. This has to be clear as we hone in on our emergent quiddity. For further purposes of clarity, I will restate my reformulations of the basic forms of metaphysical currency and how they are used in natural investigation, in order of necessity.

Figure 4.4.1: Metaphysical Currencies and their Investigation

Currency & *Example*[236]	Definition & *Description*
Property	coded sensory data
M: redness *m: reflecting red light*	• *coding requires observation* • *corresponds to—and used to describe—coded objects, events, conditions and other properties (observation-dependent)*
Object	aggregates of matter and force particles
M: apple *m: molecule*	• *individuated via properties* • *individuated by its ability to have a lesser degree of ontological and etiological dependency* than *other aggregates* on *other aggregates and the conditions they and it survive in (property-dependent)* • *existence implies suitability of conditions* • *requires instantiation for precise aggregation (property-dependent)*
Event	interaction of two or more objects
M: ripening *m: increase in sugars*	• *individuated via properties* • *event itself can generate event-level properties, which themselves can be property-dependent* • *interaction assumes instrumentally-discrete objects* • *'interaction' implies duration* • *implies possibility of multiple levels of analysis*
Condition	topographical estimate of objects and related events
M: 30°C *m: 9.81 m/s²* (g)	• *should normally imply difference in level of analysis relative to object(s)* • *often expressed in averages or ranges* • *topographical implies spatiotemporal, which implies duration*

Investigation Phase	*Definition*
Observation	*Transmission of sensory data*
Description	*Coding of sensory data into properties* *{of: objects, events, properties, conditions}*
Explanation	*Interpretation of (already-coded) properties* *{of: objects, events, properties, conditions}*
Prediction (future estimation)	*Pattern ascription to (already-interpreted and already-coded) properties* *{of: objects, events, properties, conditions}*

236 M = macro-level, m = micro-level

Because wholes cannot exist before the existence of (whatever we want to consider as) their parts, or before we decide what we are analysing, property-ascription is the first step in the construction of units of analysis. Properties lead us to—and tempt us to individuate—objects (including potential wholes and parts), events, conditions, and other properties, or properties of any of these. In this way, the *ens per alio* is the foundation upon which we can build an *ens successivum*.

There cannot be such a thing as a complete catalogue of properties for any given object. As such, the individuation of any object is only instrumental. The major implication of course is that no object can have a metaphysically-significant autonomy. But, this is not really an obstacle, because, just like when we think we are individuating a molecule or cell, the individuation itself was not metaphysically significant. We might have isolated the molecule (a difficult and expensive task) or cell from other molecules or cells, but that does not mean it is physically isolated from everything else. Autonomy, admittedly in a very rigid sense, is synonymous with ontological independence. There cannot be a sudden preference for one level of analysis in metaphysical description and explanation: 'rigid' autonomy *just is* autonomy in metaphysics, and therefore no object can have it.[237]

The object becomes an observation-driven collection of properties. We think that it has a significantly-lower degree of environmental porosity—based on 'its' properties and local events and conditions—and so we individuate it. We know it has a lower degree of porosity than other objects because it exists qua whole *simpliciter*: in order for something to have properties, it must have suitable local conditions. And, this is the most difficult aspect to verbalise. One object, which has properties, exists because it has properties, and because it does, its conditions, which are topographically mapped, are necessarily suitable for it—*quod erat demonstrandum*.

All objects require certain conditions in order to exist. Common objects, like apples and chairs, cannot exist in certain conditions, like inside a volcano. Their existence is, therefore, initially dependent on suitable local conditions. Microphysical objects also

[237] The arguments for autonomous agents a will never be metaphysically valid in a refined physicalism-empiricism. They can, however, be reformulated, for example, in a cognitive, intentionalist, or rationalist sense, to name a few.

cannot exist without certain local conditions: consider the chemical and nuclear reactions that can destabilise isotopes or bombard entire atoms. If enough energy is available, any microphysical object can be broken apart by other objects, events, and properties, which these conditions will consist of. It is the *conditions* that allow for the construction and destruction of the objects to which we ascribe properties, and this flows right into emergence.

4.4.1. Constructive and destructive emergence

In Chapter I, I showed how emergence has generally been seen in a constructive, progressive light, where a later phase or state is assumed to be more complex and/or ordered. In general, most authors optimistically envisage emergence as ‘order out of chaos’,[238] and if not necessarily with such a thermodynamic twist, emergence, for most, should at least be seen as the development of something more complex—which implies addition, and not subtraction. In general, this is acceptable, because unexpected properties are to be expected, and weak emergence is a predictable feature of inherent unpredictability.

However, with the physicalistic and nominalistic base that we have been using, objects, events, properties and conditions have no direction, nor do they have an intrinsic creative drive, nor do they have a predisposition for construction. These currencies have no master. Things exist because conditions allow them to, and not because there is an ‘arrow of constructivity.’[239]

On the surface, something that disaggregates and then re-aggregates does not seem emergent. It would take, in fact, for most authors on emergence, more than ‘aggregation’—the seemingly arbitrary and accidental collision and cohesion of particles—for something to be emergent. As we think about all of the examples in the present work of objects purported to be themselves emergent or have emergent

[238] This is also the title of Prigogine’s 1984 book.
[239] Despite modern science’s shedding of most teleological paradigms, at least the roots of emergentism were clearly influenced by variants of process philosophy, vitalism and holism—themselves in some cases dangerously close to teleology and essentialist approaches.

properties, none have passed the test of a whole's endurance. We looked at integral wholes, and *entia successiva*, and still found nothing that lived up to the emergent hype. All we had were incomplete collections of properties. How can something emerge, if we did not know what it was or is?

Emergence in its weak form helps us realise how we can easily overlook something. Emergence in its strong form, just like with ontological reduction, helps us realise how easily we can be fooled into thinking that investigation is complete. Emergence, I argue, does not require a progression to the more complex, the more dense, or the more unified. When we think we have found something emergent—like a shape, an organism, or the 'market'—a great many things had to be destroyed or be retransmitted in order for the matter and force to be aggregated as it is. It is much like biological evolution: we overwhelmingly interact with other organic things throughout our lives, to the point where we often forget the unfathomable cosmic improbability that any one of these organic things could even exist. Just because we exist, and can observe, in such a profoundly tiny biosphere, does not exempt us from seeing how the rest of the cosmos operates. Things aggregate, are destroyed, and re-aggregate countless times over without any observation or property ascription. The objects that we see, describe, and explain are there because they were allowed to be there.[240] But because they are there, *they just are*, and accounting for all causal relations is staggeringly futile.

The language of emergence is flagrantly self-restrictive. Here is a common vulgarisation: "New York City can be seen as emergent order." This dull statement expresses nothing, because we have no idea what 'New York City' is, nor are we clear how 'order', emergent or otherwise, is to be understood. Is it the excitement of the uniform yellow taxis buzzing people around marked, neatly paved streets that makes it ordered? This misapprehension is, like so many other concerns of mine in metaphysics, rooted in the conflation of the perceived and the actual. A refined physicalism and empiricism described and employed throughout this work would not claim to have unfettered access to the actual, but does know enough about the actual

[240] This statement has no bearing on a general universal claim of determinism. What brains do, for example, is not in jeopardy with this claim, as brains can do anything brains can do *under the right microphysical local conditions.*

to qualify 'emergent order' as 'merely a rough sketch of what someone claims order is.'

Every attempt to describe a macro-object as ordered, unified or even complex can be challenged. A busy city, which has no precise borders, is, like everything else, composed of matter and force, of objects, events, properties, and conditions, that appear and disappear, aggregate, disaggregate, and re-aggregate, fluctuate and evaporate. Identifying structures and agents is as difficult in a city as it is anywhere else in the cosmos. The fuel that burns in the motors of the taxis combusts, and is emitted in a hazardous cloud the drivers themselves inhale. The existence of every object in the city we nominate and ascribe properties to depends on the same rule that all aggregates adhere to: *they are there because conditions allowed them to be*.

Metaphysically-significant emergence needs to encompass whatever we consider 'construction' or 'order' as well as whatever we consider 'destruction' or 'chaos', because both streams, as vague, subjective and nominal as they may be, flow according to the same principle. Recall the example of the sand pile from §3.1. I take a random, loose grab of dry sand particles from the beach, and then dump it on a flat surface. It forms a loose, scattered pile, flattened by the particles falling over each other as gravity brings them closer to the densest centre of mass through the flat surface (the planet). The grains are separated, scattered, and some have even fallen off the flat surface. Then, the wind comes, and blows more than half of the grains away. Now, does this collection of grains sound emergent, in the ordered, constructive sense?

Imagine the same scenario, just add water. We can now take a heavy clump of sculptable sand, and, when dropping it onto the surface, we notice that it has taken the concaved shape of the hand that extracted it. Though the moisture between some particles had immediately evaporated (condition-dependent) and caused them to fall away from the main clump in the hand as it was being transferred, the majority of the extraction has remained intact and holds a general shape. Shape—one of the most basic system-level properties, which often gets the observer in the weakly emergentist[241] door—is of limited metaphysical value, and is ultimately derived from western philosophy's Euclidian and Platonic obsession with perfect forms.[242] It is a property, but our geometrical code—circles, squares, right angles—encourages us to think of the hunk of wet sand as having a discernible shape, and therefore, a unity *of some kind*. But, its vague shape does not necessitate unity: the wind comes, blows some dried particles away, causes some other moisture to evaporate, and then, with the next gust, will blow them away, too. Whether done by people on holiday or by a rare wind and tidal effect, strange sand structures aggregate and re-aggregate countless times over. There is no difference in this refined physicalist framework between the wet clump and the dry handful, and the binding of grains by water to justify unity is hardly convincing.

When a sand structure comes together, even with the guiding hand of a conscious observer, its transformation from a 'heap' to a 'whole' is not metaphysically significant. The directionality of its increase in order, cohesiveness, and unity *appears* to be singular, as the early stages of the process are clearly more like the beach as it had been found. But, the subjective, nominal, and extrinsic order that appears to increase is only with respect to whatever loose collection of particles we happen to pay attention to. The construction of a sand aggregate, which might happen to be human-made, requires the destruction of other sand aggregates. In order to build my sand castle, I need to draw sand from one general area to begin the aggregation, and, in the process, destroy whatever loose aggregate had already been there.

Emergence cannot only be one-directional, because the directionality—or increase in order or complexity—was never metaphysically significant. The central claim of *ontological* emergence is that there are exceptions, or special cases, where properties and composition break the normal physical rules. As contentious as the notion may be, it is demonstrative of this idea of irreversible creation, progress and development—an incessant superformation of ontological plateaus, as, for example, Morowitz's 28

[241] The number of times emergence is vulgarised in the name of the odd shapes of ant hills, for example, can hardly be counted. It is the equivalent of saying, "Oh, that is fun to look at. It drives my sense of natural curiosity. It is something I had not seen before, nor that I had expected to see." This is quaint, but metaphysically innocent and metaphysically meaningless.
[242] In other words, the existence of right angles, triangles, precise angular measurements—anything that pretends to be perfection—is a dream, and must be seen as useful tools for increased accuracy (but not complete accuracy).

steps (2002)[243] or Hayek's "grown order" (1973) are to be understood as. Human societies have become more complex in a weak, epistemological sense, but matter and force still compose the aggregates that evidently survive in their local conditions. Emergence must speak to the aggregation of particles, because any objects purportedly having novel properties must still be composed of matter and force (emergence presupposes physicalism, as mentioned in Chapter I, §2.1). But, since the individuation of objects is loosely based on cohesive or collaborating properties, and purported novelty appears as a result of the aggregation of specific particles under specific conditions, unpredictable and/or irreducible properties and objects are merely contingent upon convention and *épistémè* (e.g. Foucault, 1966).

Ontological and causal novelty is, as Nagel (1961) would say, a temporary admission of ignorance. If novelty is so fleeting, why assume a special place every time we have trouble describing and explaining an object? Novelty has no metaphysical or ontological opposite, because it is simply a property or object that we had not yet observed, described or explained.[244] It is by far the most striking evidence of observation-dependence, because what is mysterious and complex in one sense is simple and elementary in another. Emergence must be reversible, I argue, because the perception of order (et al) flows both ways. Aggregates with novel properties are still aggregates, and smaller aggregates with no novel properties are still aggregates. Construction from one perspective is destruction in another. There can be no direction for emergence, because emergence is blind, and there never was any direction before the claim (in the metaphysics of strong emergence) was made. An amalgamation of strong and weak emergence, then, seems to be, in its plainest form, in a refined physicalism, unexpected aggregation necessitated by condition.

[243] A curious book, it takes the reader through 28 instances of a generic, vague, macro-surprise kind of emergence, from the Big Bang, through nucleosynthesis, planetary formation, biogenesis, animals, primates, tools, language, and philosophy. Emergence as a concept is not really defined in the book, but it is a veiled attempt at proposing irreversible strong novelty (though with the attractiveness of *irreproducibility*). He uses the word 'pruning'—though having evolutionary biological connotations (e.g. Gould, SJ, 2002)—in conjunction with selection and evolutionary algorithms with regards to solutions and families of solutions.

[244] There never was a possibility of total explanation, because of ontological dependence, observation-dependence, and infinite (or rather, unfathomable) causal relata.

4.4.2. Emergent quiddities

Emergence vis-à-vis aggregation lends itself to spontaneous object and property generation manifested by suitable lower-level conditions. Particulate aggregation *is* a generic property and object emergence. Things come together and make bigger things, and the sensory data that we collect from those bigger things is coded and fed back into investigation and language.[245] In §3.4.2, I mentioned how we start with the collaborating properties of the instantiated *ens per alio* at t_1 in order to lay the groundwork for the persisting *ens successivum* between t_1 and t_3. The groundwork at t_1 is the consistence of lower-level aggregates into instrumentally-unified objects, where properties of lower-level aggregates seem to topographically converge. This nominal and parsimonious quasi-unity is acceptable, as long as we maintain our *instrumental-but-not-quite-metaphysical* caveat: in plainer terms, we can pretend atoms and molecules are discrete in order to justify the instrumental individuation of markets and cities.[246] These are objects qua units of analysis, and, in some cases, we may feel compelled to call one of them a whole *simpliciter*.

Because wholes need to have parts, and parts need to be objects, individuating units of analysis qua wholes, is, again, about reconciling, at potentially a much lower level, the property-dependent part with the part-dependent property. A loose, seemingly-dense collection of properties invites us to individuate an object, of which those properties necessarily require to be themselves instantiated. Properties cannot float, and must be attached to an object. But, the object's individuation is synchronous with the identification of a loose, seemingly-dense collection of properties. The reconciliation is done by the observer's selection of a level of analysis: I choose to make the tree branch an object, probably because it is or is part of something I can claim to see the limits of. I consolidate all of the sensory data I had coded that are loosely related to 'the tree branch' in order to ascribe properties to it. The property-

245 This is similar to Maturana and Varela's epistemological model (1988, 28).

246 We can, at t_1, establish objects, and by extension, argue for object-level properties: a benzene molecule might be said to have system-level properties, especially when we use code that is shared by other molecules—like hexagonal shape, or behaviour under conditions x and y.

dependent part has been manifested by part-dependent properties, and the part-dependent properties have been manifested by the property-dependent part. It is a mutual manifestation, and what was once an arbitrary and vague collection of coded sensory data is now an instrumentally-unified object and the *ens per alio*.

Parts, which themselves can be granted system-level properties, are constituted by objects and their properties—which we had instantiated and individuated. Some of these (constructed) parts have a discernible duration, and some do not, depending on the conditions relevant to that part. For those that persist beyond the initial instantiation to t_2 or t_3, they can be allowed to be seen as contributing to the composition of a larger, macro-object, which we have been calling the *ens successivum* and the integral whole. What is this contribution, this flow, this collaboration, this relationship between constructed parts and their meta-constructed whole?

I think this is the key to understanding the object we are studying: the unit of analysis, the thing under the microscope, the agent, the alleged emergent whole, and anything we wish to be individuated. Just as the properties of lower-level objects topographically converge on a higher-level object qua part, the properties of the parts topographically converge on—and co-locate themselves in order to instantiate—the whole qua unit of analysis. This is true for both the *ens per alio* (instantiated, or 'Line 9' in Figure 4.2) and the *ens successivum* (persistent, or 'Line 11' in Figure 4.2). A whole's quiddity is their harmony—their ideational consistence—that weakly or epistemically emerges from its physical properties. And, just like the parts, some real features of the whole will simply not be included in our unit of analysis.[247] Hence in this age of exponentially-increasing microphysical description and explanation, there will always be a discrepancy, as trivial as it may be, between the properties that we use to individuate, and the properties that we have access to. When an object emerges—or, more appropriately, whenever we observe seemingly unified aggregates—it is our static and synchronic way of capturing that which can only ever be dynamic and diachronic. *Emergence is what it claims to find*.

[247] As we have been revealing the quiddity to be ideational and non-spatiotemporal, and more of a mereological heuristic, we cross into epistemology. Perhaps this is a kind of bridge between it and ontology?

4.4.2.1. Emergence nihilism

Most of the previous sections have clearly led to a kind of instrumentality and an extrinsicality to object individuation and property ascription, and so it may, for some, paint a bleak picture of the world, or a picture seemingly not worth painting. I myself do not find it so bleak, because it strikes me as contributing to the resolution of several metaphysical issues—it seals several leaks—and yet leaves plenty of room for discussion. What should we consider an object? Perhaps a responsibly physicalist, non-ontologically reductionist approach to this question is emergence nihilism, what I think could very well become a standard assumption in future metaphysical debates.

Originally presented by J.R.G. Williams (2006),[248] emergence nihilism recognises an abundance of exotica qua simple objects that are co-located, and without proper parts. The properties we ascribe to them are extrinsic—just like the extrinsic hierarchies we might apply to seemingly organised or structured wholes. When we refer to atoms or protons, we are really just expressing (the recurrence of) the extrinsic properties observed when certain other, lower-level objects are co-located. Furthermore, there is no need for ontological primitives, a bottomed-out unit, or a fear of infinite regression (such as with the exclusively independently-partitioned *gunk*)[249]. There have been, just as there are and likely will be, infinitely many levels of analysis and smaller and smaller particles and fields being increasingly sensitive to local conditions.[250]

> To describe the actual world as the emergence-nihilist takes it to be, start by describing the actual world as a non-nihilist takes it to be. There is my body, which has arms and legs as parts, which in turn have

---

[248] It should be noted that the author does not declare outright that this is his own position, but his scrutiny of gunk and microphysical essentialism clearly leans toward mereological nihilism—and introduces emergence nihilism as an improvement upon it.

[249] Sider (1993) initially describes gunk as an atomless something that divides forever into smaller and smaller parts (and is therefore infinitely divisible). And, this sounds perfectly acceptable in light of modern physics' continued investigation and refinement of the standard model. But he also mistakenly contrasts this gunk device with a generic mereological nihilist position, where one assumes the non-necessity of mereological relata. Since the primary feature of a generic gunk is obviously appealing, and since mereological nihilism is merely guarding against hasty object individuation, I see no reason why the generic versions of these two could not play nicely together (which is, of course, a topic for another time).

[250] Even if it is foolish to suggest an 'infinite' amount of levels and units of analysis, it would be far more foolish to suggest that at any time, investigation could go no further. Again, it is very difficult to accept that, because 10^{-36}m allegedly makes no physical sense, we will never learn about anything below that scale. Without looking very far into this particular question, it would seem possible that such a momentary limit is due to the restrictions of three or even four-dimensional Earth-bound 'measurement'.

> respectively fingers and toes as parts. Ultimately, we have micro-particles such as quarks: the ultimate, simple parts of the body. Now excise from this description any mereological relations. There is my body, located in a certain place. In subregions of this place are arms and legs. In the respective sub-subregions we find fingers and toes. At the smallest [*] sub-region at which objects are located we find quarks. (2006, 504)

I would add the adjective 'observed' to "smallest sub-region", so that we might emphasise the dependence of this "smallest" label on contemporaneous physics. It would also remind us of the inseparability of object individuation from property ascription (implying observation-dependence), especially in physics. Generally I would endorse the vision Williams has for the emergence nihilist, though, unlike how this passage seems to indicate, there is probably no need to mention mereological simples (objects with no proper parts):

> The emergence-nihilist's version of the actual world will contain simples corresponding to quarks, leptons and bosons. Presumably they will instantiate the familiar microphysical properties. What of the larger simples: those shaped like atoms, molecules, dogs and cats? One option is to deny that these larger entities have anything except locational properties. They would then automatically gain certain extrinsic properties such as being co-located with two up and one down quarks, appropriately bonded. And perhaps the emergence-nihilist should put this extrinsic property forward as a candidate for what 'is a proton' expresses in the mouths of the inhabitants of the emergence-nihilism world. (2006, 505)

I find that there is no reason to supply simples, and no reason to think that any of these objects and levels of analysis have anything other than extrinsic properties. Quarks, leptons, bosons, and anything else in the standard model will instantiate the familiar microphysical properties *because we gave them to them*.[251] These properties are what our observations have yielded based on what we have individuated and how we have studied them (condition and technology-dependent). A cat has co-located parts and topographically-converged properties that we assemble, consolidate, categorise, and compose a quiddity for: "From now on, 'cat' means *X*." Nevertheless, emergence

[251] Some might object, and say, "How can you say that *we* gave electrons the spin they clearly have?" To this, I would ask how confident they are in Planck's constant. Even further: how confident are they that they are using the smallest possible unit of analysis? Why would they assume that angular momentum is not affected by other objects, events, properties, and conditions? Ultimately, the key question, though, is how any observation, and subsequent coding and interpretation of sensory data, is affected by our brains, our local conditions, and our technology. The laws (models) of physics do not really lie (Cartwright, 1983), they just blindly, unwittingly, instrumentally—though perhaps inevitably (e.g. Kuhn)—assume too much. Experimentation carries on, regardless.

nihilism is a humble and cautious way to make sense of real-world objects within a refined physicalist, nominalist and instrumentalist framework. If we take this position, it would seem that the process of object individuation, just like the object itself, is weakly emergent.

4.4.2.2. The idea of an object [252]

I have brought the concept of *quiddity* along since Chapter II because it, unlike essence, is not necessarily incompatible with our refined physicalism and empiricism. When we recognise quiddity as the ideational consistence of the *ens per alio* and the *ens successivum*, it seems to me that a physicalistic quiddity resonates quite strongly with Husserl's 'idea of an object'.[253] Whilst his *founding relation* and *founded content* both require intrinsic properties, his description of an aggregate is rather intriguing for our weakly-individuated whole *simpliciter*:

> 'Aggregate' is an expression for a categorical unity corresponding to the mere form of thought, it stands for the correlate of a certain *unity of reference* relating to all relevant objects. The objects themselves, being only held together in thought, do not succeed in forming a new content, whether taken as a group or together; no material form of association develops among them through this unity of intuition, they are possibly 'quite disconnected and intrinsically unrelated'. This is shown in the fact that the form of the aggregate is quite indifferent to its matter, i.e. it can persist in spite of wholly arbitrary variation in its comprised contents. A 'founded' content, however, depends on the specific 'nature' of its 'founding' contents: there is a pure law which renders the Genus of the 'founded'

[252] At least one feature of a Kantian legacy is easy for Quentin Meillassoux (e.g. 2008) to identify: he calls it *correlationism*, the cultural predisposition that we only ever have access to the correlation between thinking and existence, and never to either term considered apart from the other. Co-championing this idea is Graham Harman (e.g. 2002, 2009, 2010) who really seems to be concerned with devaluing post-Kantian thought and, in particular, phenomenology. One feature that seems to trouble them is the idea that no object qua object could exist observation-independently. For them, there must be something objective to which properties cling. And, as is unsurprising, Harmon actually defends or sympathizes with a neo-Platonic view of natural forms (Meillassoux is a mathematical Platonist) as well as an Aristotelian substance. He also says that *everything* is an object, including radiation, governments, and propositions. Whilst their 'anthrodecentrism' is obviously appealing to an empirically-grounded physicalism, and their object-focused perspective resonates with mine here due to objects being the most practical metaphysical currency for description and explanation, I find much of the content to be confused and pushing too far away from observation-dependence. I think parts of their cause are noble, such as what I see as the general pursuit of that which is ineligible for observational contamination and encoding. But, like so many other philosophies and posturing philosophical systems, they really just seem to be obstinately constrained by their former teachers (in this case, mainstream Continentals) in many respects.

[253] There is room here for substantial analysis of Husserl's *Experience and Judgment* (1939, 1973), but I will confine myself to *Logical Investigations* for relevance, consistency, and clarity.

> content dependent on the definitely indicated Genera of the 'founding' contents. (LI3 §23)

It seems to me that this description of 'aggregate' is precisely what I wanted to show about the reworked example of *The Ship of Theseus*. The physical contents all share a *unity of reference*, and are being held together only in thought. The form—or the system-level or macro-level unity of the object qua unit of analysis, and eventually qua whole *simpliciter*—is indifferent to variation in its physical contents. Objects, some of which being candidate wholes or potentially emergent wholes, survive compositional change (and appear integrated) because we wanted them to. The instrumental form, then, is the composite reference point, the thing for which the particles do duty, and the quiddity we use to justify individuation. In other words, if we cannot accept a genus, an essence, or a founding relation, as a refined physicalism cannot, we are left with extrinsically-organised and extrinsically-emergent objects qua wholes.[254] [255] This is weak object individuation, and so when we say that something has emerged, or that something has emergent properties, I believe that we should think of how *the idea of the object* [256] has emerged, and not of a new substance or a new category of causally-efficacious and autonomous physical systems.

4.5. Chapter Summary

Strong emergence demands that the *whole simpliciter* be individuated, and that it can have autonomy and system-level properties—at least some of which are novel (unpredictable, unexplainable, seemingly irreducible). Despite recognising the *whole*

[254] It seems to me the real beauty of this is the realisation that the process of referring all of an object's physical contents to such a quiddity—of connecting all of the detected dots to form a unity of some sort—is, on the saddle of language, and in the face of mereological uncertainty, ontologically irreducible, irreproducible, and totally unpredictable.
[255] The sceptical reader may be tempted to note that, when everything is boiled down to simple summaries, much of the content here in this present work could be described as favouring purely weak emergence—for essentially these reasons of extrinsicality. But, surely such a reader would see how we have arrived at such a place, such that when even the most robust object we can think of—such as another human being—is undeniably only *questionably* integrated, *questionably* autonomous, *questionably* emergent, it is not a trivial dismissal of the great canon of strong emergence.
[256] The emergence of this *idea of the object* is precisely what I will outline in *weak object individuation*. Whilst inspired by Husserl as already mentioned above, the method is largely independent of *Logical Investigations*: that is to say, the phrase is italicized, underlined, etc., but only to distinguish it as a concept or approach in and of itself, and not to draw specific reference to the Husserlian conception as originally introduced in Chapter II.

simpliciter's porosity—or sensitivity to and dependence on local conditions—when we put all these features together, the purportedly strongly-emergent whole qua unit of analysis still needs to persist as such an object must be observed at successive instantiations.

In order to claim that the whole has changed, we have to know which whole we are talking about: identity must precede persistence. But, this issue places us right in the middle of some of the most profound metaphysical questions, not the least of which being about what change is. With Leibniz' law making assumptions about unity and property access, and the classical conditions of change overlooking basic physical facts, persistence appears to be dependent on quiddities—and not physical composition. *The Ship of Theseus* provides us with a way to reformulate Chisholm's presentation of the Scholastic *ens successivum* whilst respecting the property-dependence of the *ens per alio*. The perdurance options appear to be limited due to the challenge of discrete time-slices, though Heller's worm perdurance shows that, like the integral whole, things wiggle through time with a kind of mereological effervescence.

Quiddities allow us to see an object as persisting, and as such they are just ideational constructs. They are the wads of yellow adhesive labels, upon which we had written properties: in order to see an object as an object, the labels have been removed and mashed together in order to make sense of an arbitrary heap. The process of quiddity-generation depends on the same topographical property convergence and co-location that manifested the parts of the very whole we wish to see as persisting. Wholes, then, persist via the emergence of quiddities when the properties of its parts and their constituent objects, events, properties and conditions converge. The ship, or any object that we often ascribe proper parts to, persists first via the overlapping of properties—vis-à-vis the *ens per alio*—which are ascribed to objects that will compose or become parts. It then persists via the overlapping of parts—vis-à-vis the *ens successivum*—which are then ascribed a role in relation to a larger, macro-object and whole.

Strongly emergent wholes, just like all wholes and objects, persist via quiddities (where quiddities are justified by persisting properties). We have to roughly individuate something for it to be autonomous, and that will always require our best guess as to the difference between a property here and a property there. Emergence nihilism provides for part and property co-location, which eventually consolidates all the coded sensory data into a quiddity. Now, we can weakly instantiate and individuate an object, and by extension, claim its persistence. As such, general emergence should be seen as a wonderful way to describe our tendency to constructively manifest persisting objects out of a swirling and ever-expanding sea of properties.

Conclusions

I will now draw the thesis to a close, in four sections: (1) how my investigation has fared considering its origins and launch; (2) how the primary questions of the work have been answered; (3) how several secondary issues might be resolved; and (4) the presentation of a final synthesis and model to complement my initial assumptions.

5.1. Expectations and assumptions

This section will address some of the basic issues going into the thesis, as well as some of the immediate problems. I will also reiterate what kinds of methodological assumptions I used, and how they helped and hindered my progress. Finally, I will discuss how some of the more basic propositions I expected to be true have held up since the Introduction.

5.1.1. Overview

As I had understood it, the formulation of claims in the metaphysics of strong emergence had been deliberately avoiding discussions of *whole* whilst at the same time requiring objects *qua whole* to scrutinise said strong claims. The discourse needed to be acquainted with the mereological, and my initial intention was to see that it became so. Furthermore, claims in the metaphysics of strong emergence, in avoiding the mereological component, had failed to see or at least capitalise on how emergence connects—and arguably contributes to—other pressing metaphysical inquiries of identity and persistence. All these problems had seemed like major omissions in what was a potential means for broad explanation in interdisciplinary challenges like structure and agency. I still consider these assessments as accurate.

After initially reviewing the core platforms in the metaphysics of strong emergence and reduction, I found that, more specifically, the discourse was littered with assumptions about not only the *individuation* of objects qua wholes, but also of their physical *integration*. It occurred to me that there was no good reason to assume that a potentially strongly-emergent whole should be necessarily individuated on the grounds that it was aprioristically integrated. The *minimum* amount of work needed on this issue was to see how wholes could be seen as both distinct and unified, in order to see if the discourse's omission of the mereological dimension was based on some standard ontological assumption.

In using a refined, non-ontologically reductive physicalism, with an awareness of modern empirical physics, delineating this individuation-integration problem unfortunately places the metaphysics of emergence directly in the crossfire between ontology and epistemology, and is what seems to me to be precisely what makes strong emergence so tumultuously digested across the social sciences and humanities. In going straight to the heart of the metaphysics of strong emergence, I believe I was justified in allowing these two traditionally distinct terrains to occasionally interact. And, considering the high stakes of strong (and ontological, I draw a distinction in Chapter I) emergence, I decided it was better to proceed with this potential confusion if I was going to try to truly make sense of emergent wholes.

What is *critical* for the reader to realise is that I have simply found it impossible to separate claims about *how a whole comes to exist* from *how we come to know wholes*. It became clear that as long as I employed this non-linear physicalism, the traditional boundary between weak (epistemological) and strong (ontological) emergence was not as clear as had been presented, and that knowing and observing wholes presented problems for what a whole is and how it comes to be—and vice-versa. Whilst perhaps it may have proven to be an obstacle at the beginning of my investigation, it helped me realise that it was probably necessary in order to make emergence *of some kind* compatible with a responsible physicalism.

With a new, non-hierarchical empiricism, I felt that I needed to have specific currencies for description and explanation. So, by using only *objects*, *events*, *properties*, and *conditions*, I tried to show, for example, that (potentially) 'emergent wholes' were objects, and (potentially) emergent properties were properties attached to objects and events (where conditions were the descriptions of lower-level objects' behaviour). This kept my analysis consistent between empirical examples, which was necessary in order to go through the various mereological options beginning in Chapter II. This approach, when combined with the refined physicalism, led much of the analysis to hold properties as exceedingly important particularly at the microphysical level. Understanding and showing the role that properties play in the designation of parts and wholes also fed back into the ontological-epistemological challenge. Properties were, in fact, their nexus, or their shared 'hub' of inquiry.

In sum, the thesis eventually needed to show how a refined physicalism and empiricism in the metaphysics of emergence—where epistemology and ontology naturally intersect—could make sense of wholes qua objects when integration and individuation seemed to be so problematic.

5.1.2. How the propositions fared

In the Introduction, I mentioned that I expected the statements below to be true, and I will now briefly address how they have fared.

- **(0)** *Candidate-objects being emergent or having emergent properties must be wholes.*

- **(1)** *If we do not know what 'whole' is, we cannot say that object S is a whole.*
    - *1a) If we cannot say that S is a whole, we cannot say that S or its property X is emergent.*

- **(2)** *If we cannot say that S is a whole, we cannot say that S persists.*
    - *2a) If we cannot say that S persists, we cannot say that S can be a whole.*
    - *2b) If we cannot say that S persists, we cannot say that emergent property X can be ascribed to S.*

5.1.2.1. Proposition (0)

The purpose of confirming Proposition (0) was explained in Chapter I. There are two main reasons why I believe it has held up. First, potentially-emergent objects needed to be objects with system-level properties, of which some were candidate-emergent properties. An emergent property is not simply a property of an object *per se*, but of a unified (in some sense), configured, organised object (possibly even of a 'system'). Second, an emergent object must also have multiple levels of analysis, where the higher level of physical reality is 'doing something it should not be doing' based on knowledge of the lower level. Hence, objects contained within the system or object are therefore *parts* or *contents*. Concisely then, any claim about a potentially-emergent object is a claim about a potentially-emergent whole, because the claims are necessarily about the relationship between what we know about the higher system-level and lower level objects. These are claims that are, at least initially, before mereological options are selected, about individuated and integrated objects—which we can also call *wholes*. Proposition (0) holds true when using 'object' in the spatiotemporal way that I have.

5.1.2.2. Proposition (1)

Proposition (1) on the surface indicates that the metaphysics of emergence cannot proceed until a mereological option is chosen. I think that because we are left to assume that the potentially-emergent objects are indeed aprioristically integrated and individuated, the discourse hopes to proceed with its analysis of macro-level causal powers (for example) assuming the candidate object is indeed a unified agent. I have argued that (1) is true, because in a responsible physicalist framework, no level of analysis is excluded in the scrutiny of the claim regardless of its microphysical volatility. I have also argued that the dependence of such a framework on contemporaneous physical models prevents it from being a closed question: whatever standard model we have today may not be what we have tomorrow. Therefore, in conjunction with the already-given post-Copenhagen microphysical indeterminism, no candidate object on any level of analysis could possibly be seen as a closed and metaphysically-autonomous system. Judgment must be suspended in claims for ontological novelty as well as for reflexive downward causation (for example).

Secondly, because we have been assuming that properties can only be attached to objects (and other events and conditions, within which objects behave), and because emergent properties are still properties ascribed to wholes, it seems unlikely that an emergent property could be ascribed to an object whose composition was undefined. Though the higher level is indeed doing something it should not be doing based on knowledge of the lower level, we should withhold strong claims until we have reason to see the object as integrated and justifiably individuated in a metaphysically-significant sense. From my perspective, using a refined physicalism and empiricism, Propositions (1) and (1a) hold true.

5.1.2.3. Proposition (2)

With regards to Proposition (2), if we cannot be certain that S is a (metaphysically-integrated and individuated) whole, and/or that S's physical composition is only partially known at t_1, then we cannot compare our knowledge of S's composition at t_1 with our knowledge of S's composition at t_2. Both would necessarily be incomplete, and so the only way for them to be compatible, and to see them as possibly persisting in some way, is either to exclude certain levels of analysis (and/or various microphysical objects or their classes) or to specify the nature of this potentially-emergent object's endurance (where it wholly persists).

Conversely, Propositions (2a) and (2b) point out that we cannot ascribe emergent properties to that which does not survive between instantiations. In the most basic sense, that which only exists at t_1 cannot be an object. We would have no reason to consider S as even remotely worth individuating (let alone integrated) if it popped in and out of existence with only one observation. Objects must therefore have a duration greater than zero, which can then be divided infinitely. Objects must persist, and since a potentially-emergent whole is an object, that which does not persist cannot be an emergent whole or have emergent properties. These claims bring us full circle to the requirement that an emergent whole be an integrated and justifiably-individuated object, and that when using the admittedly-stringent framework that we have been using, claims about autonomous macro-level causation and even property-ontological irreducibility must be suspended.

5.1.3. Unintended outcomes

One of the more surprising elements of my investigation was the contemporary usefulness of the Scholastic inventions of *totum integrale* and *quidditas*. The former had been devised as a third category of 'whole' to accommodate objects (and their classes or topics) whose composition did not wholly persist. The latter term was a conventional distinction to an Aristotelian 'essence' that recognised the need for the philosopher to see what an object really is without it necessarily possessing a substance and therefore an essence from which we derive the nature of the given object. Both of them proved to be critical concepts all the way through to the development of *weak object individuation* in Chapter IV. Another interesting surprise was the unexpected relevance of Husserl's 'idea of the object', which lends itself to the view that there need not be a dichotomy between the spatiotemporal and the non-spatiotemporal. This is of course a massive area of inquiry, beyond the scope of this work, but I consider the prospects of this line of thought—which admittedly is more appealing for me in the context of his discussion on *heaps*, and not on *founded contents*—to be rather strong in developing my ideas in Chapter IV more thoroughly in future projects.

5.2. The primary issues

I will present here in this first section, hopefully with some concision, how the major questions of this work have been answered. This will allow the individual Chapter macro-questions to be seen as addressed and (hopefully) relatively resolved.

5.2.1. Does emergence need a whole?

Chapter I emphatically shows that claims in the metaphysics of strong emergence necessarily refer to wholes, whether it be in the name of system, object, unit, or higher level, because strong claims depend on unified agency. As discussed in §1.2.1 above, all of these synonyms have constituents, and therefore all have constituent parts, or objects, events, properties and conditions.

Weak claims, by contrast, do not require unified wholes, because they reflect an irregularity or inconsistency (or simply general confusion) in the coded and interpreted sensory data. They can and should be ubiquitous, pan-disciplinary, instrumental, and metaphysically innocent. They are and have always been an essential part of any investigation. Every time an observation yields strange, seemingly new, unpredicted properties, we learn more about the natural—and, by extension—the social world. The position behind strong claims presupposes adequate knowledge of the phenomenon in question, whilst weak claims avoid making any such statement. They can reflect a willingness to discover and to challenge existing models, and they offer the chance to increase one's awareness of things around them, past and present. They are purely exploratory, and therefore helpful. Weak emergence and the lessons of epistemic novelty can evidently be applied to quiddity construction and eventually to weak object individuation.

5.2.2. What is a whole, and how can it be determined?

Classically and conventionally speaking, a whole is that from which nothing or no part is missing.[257] This presupposes the knowledge and awareness of totality, completeness, symmetry, and often perfection. The strong claims of both emergence and reduction—and ultimately all of their analysis—grant such completeness, and, in some cases, invoke the non-spatiotemporal to explain the unity that such a totality

[257] The importance of this distinction was first mentioned in Chapter II §1.2, initially presented as potentially a key to classical mereology. As we have progressed, I have tried to show (explicitly at the beginning of Chapter IV, and implicitly throughout Chapters III and IV) why such a distinction between these has become irrelevant in the present context.

would require. Whether it be form, substance, essence, soul, configurational force, or mathematical unity, something *real* was needed to make 'the all' into 'the one', thereby making the object supersummative: the state of being more than what one knew about its physical components.

But, what *does* one know about an object's physical composition and properties? Such knowledge is, evidently, increasing—or, at least, changing. This is the nature of investigation. As it turns out, though, strong and supersummative claims, of both types, throughout the history of philosophy, have depended on physical *assumptions*: enough is known about *X* such that one is compelled to support claims of compositional unity, structural organisation, and autonomy. The ontological reductionists have their primitives, from which descriptions and explanations look upward from said primitive. The strong and ontological emergentists have their novel unit of analysis, from which new causal powers look inward, downward, and outward into the world often inspiring claims of new kinds of entities. Unified wholes—and unity in general—comes, for these groups, from either the bottom-up or from the top-down.

5.2.3. How can a whole change?

Of all the questions that the metaphysics of strong emergence does not address, this is perhaps the most embarrassing. If there is a difference between a heap of sand and an organism, with regards to organisation, autonomy, and unity, how do these unified entities persist from one moment to the next? How do they survive change? The physical wholes used in the debates in the metaphysics of emergence and reduction are made of physical things, whether or not we are including supersummative elements. Regardless of the omission of *any* discussion of the status of their unity, physical wholes simply do not persist in a way that would fit a unified entity. Whether or not one wants to recognise parts, properties, or both, the objects to which strong features have been ascribed *unambiguously fail* to sustain themselves—and we do not need to delve into quantum foam to demonstrate it. Since all wholes are susceptible to change, if only at the microphysical level, then a whole that persists is a whole that has lost some parts and properties. The *totum integrale*, as it was originally formulated, still has an essence that determines which parts are not needed to persist. That leaves us with the whole *simpliciter*, which, after seeing it as the consistence of the *ens successivum* and the *ens per alio*, manifests a conventional quiddity that binds the overlapping contents.

5.3. Analysis of secondary issues

As an outgrowth of the original intention of assessing strong emergence in the context of wholes *per se*, and as a consequence of connecting emergence with a mereologically-conscious refined physicalism and empiricism, I believe several other statements and possible alternative conclusions can be made.

5.3.1. Claims about both standard reductive physicalism and strong or ontological emergence actually make the same mistake

As I have tried to emphasise throughout this work, standard physicalism's association with ontological reduction is a very difficult one to shake. Considering that it relies on a bottom-level or some kind of unobserved ontological primitive (e.g. string), it is making the same hasty individuation as the autonomy of higher-level units of analysis in strong emergence. The only discernible methodological difference between them is vague at best. For the former camp, our minds have little trouble visualising the causal and compositional reduction of a clock, but struggle with the atom. Therefore, as the argument goes, by the time we have reduced-away to the level of the quark, no one will understand this low level well enough to question the importance of vertical reduction. The strong reductionist vis-à-vis standard physicalism argues that the ultimate bottom level (which is real) is some member(s) of the standard model or its (their) aggregate, because causal explanation at such a low level is seemingly more linear, less complex, and ultimately predictable. The strong emergentist, initially fooled by this in accepting ontological reductionism *a priori*—and being content with the reduction of the clock but not of the organism—claims autonomy of the higher-level because the linear and vertical explanations that had suited the reductionist's strategy of convenience do not suffice. The end result of this is both sides not realising they are playing the game according to the same set of wrong rules.

It is my belief that the inability of either side to see the instrumentality of object and quiddity ascription in order to justify individuation is really based on an assumption of *integration*, and that assumption of integration is either an oblivious one or an instrumental one. When a level of analysis is excluded when an object is assumed to be an individual, such an omission could only be explained, in the first instance, by a lack of awareness of the physical reality relevant to the object. Is it not apparent that a complex system, such as an organism, does not sustain its physical composition (e.g. 98% of the average human body's cells being replenished within three months)? At the microphysical level, why would any particular level or unit of analysis be exempt from conditional dependence, and when recognised, do these conditions not necessitate respectively higher-level phenomena (e.g. the creation of heavier elements from intense atomic fusion)? In the second instance, if the decision to ignore compositional fluctuations was for investigatory and practical reasons, why bother making strong and even radical claims for isolated autonomy, an ultimate bottom level, or new categories of existence? It seems to me that both extremes are really just saying the same thing, with each side picking their own preferred level of analysis. I would also add that, though it is outside the scope of this work, the key question may very well be not about the level or unit of analysis from which agency is necessitated, but rather about how agency is observationally manifested by concentrated *powers* (which we identify by coded sensory data—properties) at any designated level of analysis. Hence, this line of inquiry could very well be in the history and philosophy of science, and not in metaphysics, ontology, or even epistemology.

5.3.2. A refined non-reductive physicalism and empiricism denies the possibility of summation

Perhaps the primary reason for modifying the physicalist position—thereby creating what I have been calling a refined physicalism—is the need to accept one single physical fact: no system can be closed. No system can be observed that does not exchange matter and force with its environment. The implications of this fact, though, are perhaps not very well understood or have not been applied to the discourse—and probably several areas of metaphysics. I have consequently taken it upon myself to demonstrate the significance of this for object individuation. In the introduction, I posed the question of how something could be supersummative—having powers or properties existing outside of an object's known physical limits. The word itself implies a knowledge of *sum*: the totality of parts, powers, and properties. If all objects are open, and therefore porous, there must be uncertainty about a summation of those parts, powers, and properties—particularly across observations. Investigation's need for individuation must, then, come to terms with an imperfect totality and an imperfect sum. She who would become a strong emergentist, as she is required to accept (at least initially) exclusively physical composition, must therefore acknowledge that any forthcoming claim about a strongly emergent object is necessarily dependent on, paradoxically, an inexorably-incomplete whole. This is at the core of the thesis.

A claim for—or about—supersummativity, we should say, must be a *weak* claim and be based on a consciously-incomplete catalogue of properties. The dependence of classical mereology, and much of the emergentist and reductionist literature, on forms, essences, configurational forces, metaphysical atoms, ontological primitives and sets—being examples of things that have engendered wholes—has always been induced by the illusion or instrumentality of unity. This unification allows the whole to be individuated and organised, whilst providing for persistence of some kind—as I have tried to reconstruct with the instrumental *quiddity*.

5.3.3. The emergence worldview (weak or strong) unwittingly illuminates local condition-dependence

Whether in weak or strong forms, claims of novelty show us that investigation is on-going, and more importantly, that the properties of our units of analysis—derived from the objects, events, properties, and conditions they depend on—are, like the observer, subject to local conditions. Some properties simply last longer than others, as conditions allow. We observe only what survives—and never all of it. When we observe something strange, we already know that the conditions—which we could never know assiduously—must have provided for it. The enabling conditions might be on the same level of analysis as our novel property, exclusively lower-level or even microphysical conditions, or possibly a combination of both. New ontological categories and new *de facto* causal powers do not emerge, but new properties do—and those properties help us decide which aggregates are objects and whether they have powers. I believe this is the key to emergence and the reason why emergence matters.

I say 'unwittingly' because throughout the various ebbs and flows of the extreme positions on both sides, going back over a hundred and fifty years or so, ontological dependence has never played a central role in the debate. For every morsel of truth 'captured', whether it be the uniqueness of the organism or mind, or the fundamentality of the atom or quark, it was never proposed that novelty's source was the unknown itself—and not the unique, individuated object. Elucidating the significance of this idea—which is of course a shared idea, and a synthesis of many different speculations and observations in various disciplines—is a key area of future investigation for me.

Early in the work I mentioned how *emergence* for many must entail *event* and non-instantiation through the assumption of what could be called an ontological diachronic emergence. I hope I have adequately shown how *my* events and conditions (which include events or even Whiteheadian processes) have incorporated this dynamism into the ascription of properties to objects. This volatility is precisely why I have tried to demonstrate why object individuation is both dependent on, and vulnerable to, a mereology and ultimately a quiddity. How we arrive at that quiddity—and the role of the idea of the object—is what I hope makes weak object individuation so interesting and useful.

5.3.4. A refined non-reductive physicalism and empiricism seems to lead to mereological nihilism

We have seen how the illusion of unity leads to the construction of quiddities and wholes (and their parts). At first glance, our framework leaves us with no objects *per se*, let alone wholes *per se* or *entia per se*. In this respect, the prospects of making sense of real-world objects, structures, and agents, seem rather bleak, because if nothing is metaphysically unified, and all we have seen is metaphysical disjuncture and condition-dependency, one immediate question is: how far could our causal explanations go? There appear to be no proper parts, no unified wholes, no autonomous agency, and limitless or at least unfathomable environmental influence. The part-whole nihilism that we are left with seems to expunge much if not all of the meaning and optimism of our observations, descriptions, and explanations, despite the fact that this nihilism has developed from such investigation.

Are we to consider such a physicalism as profoundly *unhelpful*? Does it matter if something is unhelpful, or if it makes it more difficult to investigate or philosophise? Does mereological nihilism have anything to say about the real and the non-real? For example, if we believe that an organism is metaphysically non-unified, has no essence, has only arbitrarily designated parts, and persists as an *ens per alio*, should the organism even be considered as 'real'? Are there different levels of truths about real objects, where an organism is not real because it is necessarily an *ens per alio* empirically-speaking? Or, is an organism real because it is an agent in the context of 'quidditious' powers? In other words, should all criteria for 'real objects' be seen as open-ended models and purely speculative philosophy, remaining fully compatible with, and resting firmly upon, a foundation of assumed irreproducibility and fundamental disunity, non-integration, and necessarily-instrumental individuation?

Or, are we back to the classical question of whether mental events are real or not? Husserl's 'idea of the object' could be relevant here, but perhaps the problem is with the general consequences of *any* physicalism—where all observations are necessarily of physical things. I suspect, however, that it is merely a question of objects and powers, and how we use our coded sensory data to individuate and ascribe them. The belief that the physical world is indifferent, and without perspective, and that it has not provided and cannot provide a canvas from which we derive universal truths, should be, in my estimation, a rather encouraging and open-ended thought.

5.4. Synthesis: *Weak Object Individuation*

We know that investigation requires objects to be individuated—in some way, on some grounds—in order to describe and explain. So what can be done with the refined physicalist's acknowledgment of unavoidable disunity and compositional porosity? On what grounds, or by what means, can we salvage an object qua unit of analysis? Is there any room, in this approach, for objects qua wholes? Before we have wholes, we would need objects; before we have objects, as I have argued, we would need to start with properties.

I have gone to great lengths to show that, within this new physicalism and empiricism, real-world examples of uniform shapes, smooth objects, mathematical precision, metaphysical integration, and compositional unity, must be seen as a product of level-exclusivity (and in some cases, a product of mental events). Spatiotemporal objects, as we look closely at them, are messy and unbound, much like a garden shrub. As much as one tries to control its shape, there is always a branch or two that is slightly out of place or has a tendency to grow differently at contrasting speeds, thicknesses, and angles than the others. The professional landscaper's best efforts will probably produce a stunning symmetry from a distance, but closer inspection reveals several rough edges and distinct imbalances. Our description (and eventual behavioural explanation) of the object is dependent on the observer's level of analysis. The lower the level, the more properties we have access to; the more properties we have access to, the clearer the picture. I must re-emphasise that this is not to say that studying lower-level *objects* necessarily improves our understanding of higher-level objects, events, properties or conditions. It does mean, however, that the larger the catalogue of properties—the more dense our picture is—the better our observations will be. This means that whilst pure ontological reduction is untenable in this framework, property reduction is not—when we assume that properties are merely coded sensory data (as I have been assuming in this work). When we stand back, and prefer to gaze at the patterns and designs of the garden from afar, rejecting the knowledge of the imbalances and imprecision within the shrub, I believe we are doing precisely that which we do when we individuate physical objects.

The decision to individuate the objectively-fuzzy object may derive from ignorance (as we have with perceptibly-integrated organisms for thousands of years), momentary self-indulgence (as we do with pieces of art, for example), disciplinary focus and parsimony (a given genetic research project will not account for quantum fluctuations), or a philosophical instrumentality (it was not necessary to include a given level of analysis to prove a particular point). Ultimately though, we need to, anyway, whether or not it is convenient or relevant to the claim at hand—as it is for us now. Because investigation and language require it, our physicalism tells us what objects are made of, and our empiricism tells how we can study it. The objects of any investigation being necessarily physical, whether they become wholes with parts, or agents with powers, are known by properties.

What follows in this section is the process by which I believe a refined physicalism and empiricism could hope to justify object individuation: I refer to the process as *weak object individuation*. There are several steps, with each one building upon the previous. Observation—the transmission of sensory data—is the first step in the acquisition of properties. From this transmission, our brains begin filling in the picture where objects can be distinguished, agents can be identified, and wholes can be constructed.

5.4.1. Local physical conditions manifest observable aggregates

This is a necessary first step, considering the foundational importance of ontological and conditional dependence. Physicalism means that all objects are necessarily physical, and together with empiricism it also means in this case that the objects we observe (and whose behaviour we try to explain) are therefore made of physical objects, events, properties, and conditions. They also mean that the observer is conscious and capable of detection (and computation, if applicable) of sensory data because local physical conditions permitted it. So, both the object of perception as well as the observer are dependent on suitable local physical conditions. I believe it is important to realise that local physical conditions not only *permit* the existence of both the object and the observer, but they also completely determine and control the nature and the transmission of the sensory data that will eventually become coded into properties. The example of the apple from Chapter IV might help here: Person 1 in a lit room says that an object A is a red apple, whilst Person 2 in a dark room says that A is small black ball of some sort. We have no grounds for saying that A has any intrinsic properties, and that is because the observer depends on local conditions to assimilate their sensory data into properties. Therefore, in sum, local physical conditions permit the aggregation of matter and force particles, and local physical conditions alone permit the detection of such an aggregate by an observer.

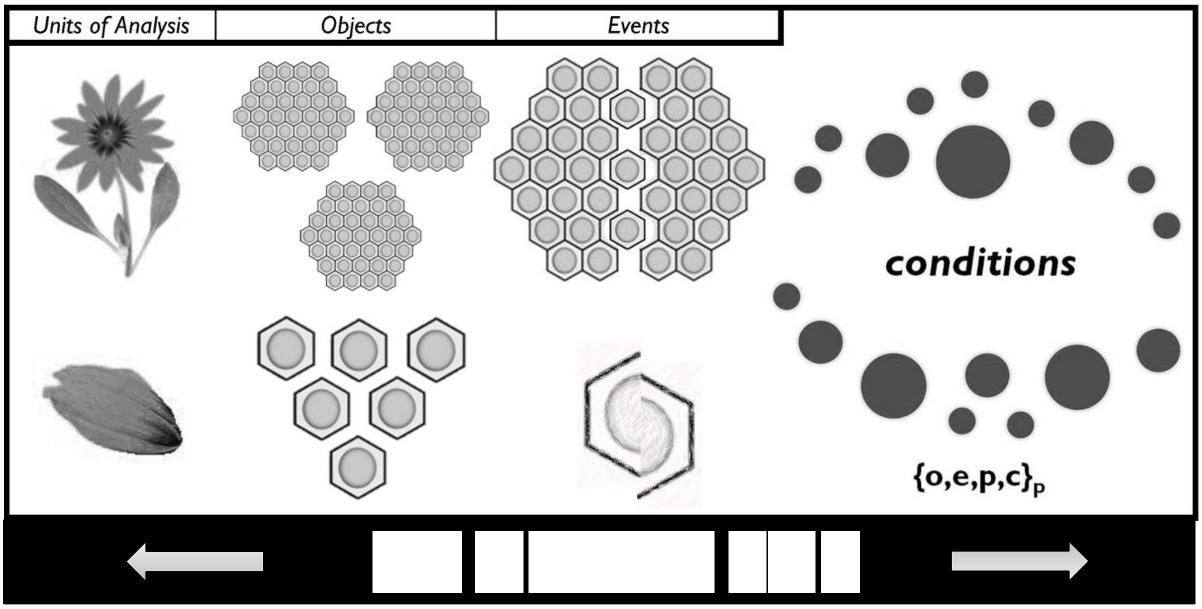
5.4.2. Topographical property convergence enables weak micro-object individuation

Properties topographically converge when they originate from a perceptibly-shared region of space-time. The yellowness of a flower's petal is derived from the perceived yellowness of many seemingly-contiguous, arbitrary regions of the petal. If we were to look under the microscope, we would find that large stretches of the petal were indeed yellow, but by no means was the entire petal yellow—as there were small brown and green micro-regions of the petal our first glance did not detect. If we look less closely, and perhaps more routinely, 'this region of the petal here' has a similar yellowness as 'that region of the petal there'. The yellowness of these and other regions of the same petal allows us to consider the entire petal as having the property 'yellow'. In the *likely* event that other contiguous petals of a given flower also share a similar yellowness, or in the *unlikely* event that we examine all petals in this same manner, we may eventually conclude that the flower has the property 'yellow'. Thus, in a refined physicalism and empiricism, the ascription of a property to an object inevitably, though not necessarily consciously, begins with lower-level properties converging on the same region of space-time. Since we have been assuming that objects are aggregates of matter and force, with which we can associate properties, and to which we can ascribe them, we could begin to see this this kind of convergence as being potentially applied to all aggregates qua objects when such observation prompts description. This is, for me, evidently a necessary step in beginning to understand object individuation.

5.4.3. Weak micro-object individuation enables weakly-persisting micro-object co-location

Once we have decided to associate the converged properties of aggregates with a singular object, we afford ourselves the ability to reinforce this weak individuation in our own minds. We do this by claiming that, because these properties may manifest themselves at successive instantiations, our particular object—our hub of properties—appears to persist from t_1 to t_2. It seems to be the same petal as before because a) my observations (sets of sensory data) appear the same or similar, and b) all of the properties I ascribed a moment earlier are still able to be ascribed. Therefore, 'the petal' has weakly persisted, and it was advantageous of me to ascribe all of those properties to it. Furthermore, now that I think about it, the other adjacent petals are quite similar to the one I had observed, and it is conceivable that they belong to some group and have a special relation to something bigger.

Figure 5.4.3.1: Weak object individuation



5.4.4. Weakly-persisting micro-object co-location enables weak macro-object individuation

If we agree to say that 'the petal' is an object, because such individuation can accommodate the persistence of the converged properties, we can often find other examples of other objects with similar properties that also persist. At this point, we may create a class or kind of objects having these properties, despite the fact that we know that a) the only reason we individuated it was because we had enough properties converging on a roughly-designated region of space-time, and b) the object seemed to sustain *enough* properties for our two observations to agree. Making these assumptions of weak persistence of converged properties, we can claim that multiple petals, belonging to a class or kind, or at least sharing *enough* properties, stand together, co-located, where they compose a further object. There appear to be several objects that we have been calling 'petals' that have survived to a later observation, and they are co-located in such a way that when seen together—like the regions of the petal—they, as Chisholm would say, do duty for a higher-level object. This is weak macro-object individuation, the *raison d'être* of emergence nihilism.

5.4.5. Weak macro-object individuation enables weak macro-object persistence

The petals, weakly individuated via topographical property convergence, are co-located in the larger area of what we would now call 'flower'. Whilst biological investigation tells us that some description of the flower's development comes via an understanding of 'organism'—vis-à-vis the properties $\{n_1 \ldots n_n\}$ associated with the flower qua object *F* being to some degree dependent on a genetic algorithm—we are nonetheless trying to see the example in terms of real-world objects. *F* has weakly-individuated objects qua parts, like the petals, the leaves, the roots, and the stem, but the process of such individuation is the same for all of these things. Referring to these parts $\{P_1 \ldots P_n\}$ as *F*, furthermore, does not preclude any strongly or metaphysically significant distinction between them or with any non-F. In other words, with this new physicalism, just because we say that the flower is made of these parts, does not mean that we are really saying anything significant about the flower or its 'parts'. It was all just an approximation based on arbitrary object and property co-location, and our individuation reveals itself to be even weaker when we try to see objects as persisting.

Another means of reinforcing and reassuring our decision to weakly individuate the micro-object is seeing the macro-object as persisting in a similar way. If the petals persist from t_1 to t_2, especially along with the leaves, roots and stem, does whatever they appear to compose not also weakly persist? The converging properties that enabled our micro-object individuation either persist or re-manifest themselves (due to suitable conditions), whilst the petals, leaves, roots, and stem that we had weakly individuated are co-located at both observations. We can see that, in this example, we might feel safer calling these things ‘parts’ because they have a well-documented connection to each other (e.g. a nutrient delivery ‘system’)—which we can observe when, for example, we separate the roots from the stem, with an ensuing change in the properties of the petals. But, in the context of the flower qua object, our knowledge of *F* is still limited to the sum of the properties of whatever micro-objects we had weakly individuated. The existence and persistence of the weakly-individuated flower (*F*) is to some extent dependent on the existence, co-location and persistence of its weakly-individuated parts $\{P_1 \ldots P_n\}$, which in turn are to some extent dependent on its topographically-converged properties $\{n_1 \ldots n_n\}$. To compound and confuse this even further, these conceptions feed back into each other, creating linguistic conveniences to make sense of the dynamic aggregates qua real world objects.

5.4.6. Weak macro-object persistence enables macro-object quiddity generation

The weakly-individuated macro-object at both t_1 and t_2, composed of seemingly-persisting weakly-individuated parts at both t_1 and t_2, tends to develop its own character the closer its level of analysis is to our own human, individual level. A flower is something we can hold in our hand, unlike that of an isolated molecule or large landmass. Because of this scalar proximity, we have an easier time turning objects like flowers into kinds and categories, especially when they are in such abundance. There are many yellow flowers in the biosphere, with many different genotypes and phenotypes, of which one might choose to organise. In a non-linear, non-hierarchical, irreproducible physicalism and empiricism, however, this organisation is mere instrumentality: a helpful, efficient means of grasping the diversity of plant life on this planet. For the purposes of language and science, we have come to see kinds like ‘flower’ determine, or at least permit, the properties a given member of that kind is able to have. Perhaps in some cases inadvertently, by classifying objects in this way we are assuming our own property-ascription is sufficient—and this is never, without exception, the case. Pretending to know the intrinsic *what-ness* of *F*, such as its essence, in reference to anything other than honest instrumentality, is metaphysically and intellectually crippling, because this is precisely the inverse of how objects become individuated: designated properties do not run to the nearest object! They are ascribed because our observations prompt them, and they will always be non-complete, non-contiguous, and most importantly, non-unified.

It is this inversion, I believe, that paves the way for definitions of perfection, completeness, unity, normalcy, and, eventually, rules of behaviour. This is the construction of quiddity: a seemingly harmless, seemingly necessary injection of a qualitative standard into general investigation. The frustrating assumption that we might know *what an object is*, that we might know *enough* about it relative to no purpose, that we might have access to intrinsic properties of any kind, these are all frightening and archaic notions. And, in my opinion, these notions are at the root of many of the problems of not just the existing emergence literature, but it might be said that they are found in several areas in metaphysics. 'What is a flower?' is an interesting question, not just because it will prompt complex answers from biology or physiology, but because it leads us down the path of how we came to know *what a flower might be*. The response that 'a flower has petals, leaves, roots, and a stem'—whilst begging for infinite definitional regression—immediately and almost punitively reminds us that it is in fact properties that determine what a flower, and any object, is. As soon as we embrace that, and the focus of investigation is property-based and not unity-based, we see our units of analysis as they are: aggregates with strange properties.

5.4.7. Quiddity generation casts the illusion of unified persistence

When we say that studying flowers gives us knowledge of flowers, we mean that many observations of *things-like-F* have enough of the same properties to create and assign to their own category or kind. An object's quiddity is therefore dependent on, and directly and exclusively informed by, membership in such a category. For every predicate 'is an *X*', we acknowledge the non-spatiotemporal existence of *X* as determining the object's quiddity. If *F* is an *X*, then we know F's quiddity, and therefore have grounds to assume integration and justify individuation. From there, we could then claim to know enough about *F* to potentially determine its intrinsic properties, its functions, its capacities, possibly even its goals or ends, and in the context of persistence, we could even determine what parts it needs to survive physical change.

Throughout this work, I have asserted that an object's quiddity determines what needs to be present at any given instantiation for it to be that thing. Its membership in a given class, category, or topic, or, more commonly, its definition, all dictate the terms of its survival of physical change. If a tree is cut down in the forest, leaving only a stump, it ceases to be a tree. If a plane crashes, leaving only loose debris, it ceases to be a plane. Once we agree to a quiddity, usually through definition and convention, we establish existence and non-existence of F. If the quiddity of F is level-exclusive (as presumably all quiddities of physical objects would need to be), F can survive inevitable physical change on the excluded levels. Moreover, the F that survives physical change is the F that remains unaffected by those excluded levels. Without this new framework, and without seeing F's quiddity as *emerging* from the collaboration of all the preceding steps of this process of object individuation, we would believe that F survived because it was integrated across observations. It must be clear, I argue, that an object's what-ness is directly responsible for supplying meaning for investigation and language, despite the fact that determining this what-ness all began as scattered property ascription. This is our emergent *idea of the object*. This is our emergent whole.

5.4.8. Final thoughts

The major framework used in this work—what I have been calling a refined physicalism (and empiricism)—is indeed what makes this entire investigation possible. It is what allows us to see cloudy hunks of matter and force particles-and-fields as observation-independent. They are somewhere, they can be observed somehow, and our consent is not required. How we individuate them, how we know they are integrated, how we ascribe properties to them, how we track them, how we explain their existence…these are all secondary or tertiary issues. Just as I have said from the very beginning, the metaphysics of emergence needed to start with *something*: a physicalism rooted in physics. From there, I believe we have been able to justify the limiting of our discussion to these four types of currency, and I believe that it is because of this that we have been able to cover so much terrain in an impossibly-vast arena.

Emergence, initially cast as an exception—a red flag—for normal description and explanation, I believe should actually be seen as the rule for object individuation. The key implication of this notion unexplored in this thesis is that of the (weak) emergence of *powers*—though such a discussion would also begin necessarily with properties (or properties qua powers). As I have said earlier, emergence is what it seeks to find. We determine consciously or not which aggregates are objects, we determine what objects to study, and we determine the coding of properties required by description and eventual explanation. Intentional *and* not, wholes are our creations that bounce between our constructed quiddities and future observations.

Additionally, if agents and structures are intended to be wholes in virtue of their distinct and opposing autonomy, and wholes are useful, illusory constructs of level-exclusive unity, is it not obvious that human agency and whatever selective structural forces one might feel threatened by are as fleeting an aggregate as a heap of sand? What argument could overturn such a statement, assuming we still use the same new physicalism and empiricism? It seems to me that if it indeed leads to the understanding of human agency and personhood as such a construction, and that all properties of human individuals and their behaviour can be reduced to weakly-individuated physical objects, events, properties, and conditions on lower levels of analysis, the framework could be seen by some as unconvincing because it makes philosophy and social inquiry entirely dependent on the physical. But, since this is precisely what is to be expected from a physicalist presupposition, and considering the weaker mereological alternatives, it does perhaps present us with an opportunity to reimagine the emergence of powers in a metaphysically-innocent way.

References

- Alexander, S (1920) *Space, Time, and Deity* Macmillan
- Archer, M (1995) *Realist Social Theory: The Morphogenetic Approach* Cambridge University Press
- Aristotle (1924) *Metaphysics*, trans. WD Ross, Clarendon Press
- Aristotle (1936) *Physics*, trans. WD Ross, Clarendon Press
- Assad, A.M. and N.H. Packard (1992) "Emergent colonization in an artificial ecology" in Varela, F.J. and Bourgine, P., eds, (1992) *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life* MIT Press 143–152. Excerpt of article also found in Bedau et al (2008)
- Augustine (1998) *Confessions* trans. H Chadwick, Oxford World Classics: Oxford University Press
- Bailey, C (1964) *The Greek Atomists and Epicurus* Russell & Russell
- Bain, A (1870) *Logic Books II and III* London: Longmans, Green, Reader & Dyer
- Balashov, Y (1999) "Relativistic objects" Noûs 33 644-662
- Bedau, M (1997) "Weak Emergence" Noûs 31 *Supplement: Philosophical Perspectives* 11 Mind, Causation, and World 375-399
- Bedau, M (2008) "Downward Causation and autonomy in weak emergence" found in Bedau et al (2008)
- Bedau, M and Emily C. Parke, eds (2009) *The Ethics of Protocells: Moral and Social Implications of Creating Life in the Laboratory*, MIT Press, Cambridge, MA
- Bedau, M and P Humphreys (eds) (2008) *Emergence: Contemporary Readings in Philosophy and Science* MIT Press
- Berger, PL and T Luckmann (1967) *The Social Construction of Reality: A Treatise in the Sociology of Knowledge* Anchor Books
- Bergman, Gustav (1944) "Holism, Historicism, and Emergence", Philosophy of Science 11 209-221
- Bhaskar, R (1998) *A Realist Theory of Science* Verso Classics
- Bidwell, R G S (1979) *Plant Physiology* Collier/MacMillan
- Black, M (1952) "The identity of indiscernibles" Mind 61 153-64
- Bogaard, PA (1979) "Heaps or Wholes: Aristotle's Explanation of Compound Bodies" *Isis* 70 No. 1 (Mar 1979) 11-29
- Bourgine, P and J Stewart (2004) "Autopoiesis and cognition" Artificial Life 10 327–345
- Boyd, R (1980), 'Materialism with reductionism: what physicalism does not entail', in Readings in Philosophy of Psychology Vol 1 ed. N. Block, Harvard University Press

- Broad, CD (1925) *The Mind and Its Place in Nature* Routledge & Kegan Paul
- Buchdahl, G (1969) *Metaphysics and the Philosophy of Science: Classical Origins Descartes to Kant*, Basil Blackwell
- Burkhardt, H and W Degen (1990) "Mereology in Leibniz's logic and philosophy" *Topoi* 9 1
- Cartwright, N (1983) *How the Laws of Physics Lie* Oxford University Press
- Cartwright, N (1989) *Nature's Capacities and Their Measurement* Oxford University Press
- Cartwright, N (1999) *The Dappled World: A Study of the Boundaries of Science* Cambridge University Press
- Cartwright, R (1975) "Scattered objects" in Lehrer, K (ed.) *Analysis and Metaphysics* Dordrecht
- CERN (2011) "Measurement of the neutrino velocity with the OPERA detector in the CNGS beam" Arxiv: 22 Sept 2011: http://arxiv.org/abs/1109.4897v4 Accessed: 18 Feb 2013
- CERN (2012) "Measurement of the neutrino velocity with the ICARUS detector at the CNGS beam" Arxiv: 29 Mar 2012 http://arxiv.org/abs/1203.3433v3 Accessed: 18 Feb 2013
- Chisholm, R (1967) "Identity through Possible Worlds: Some Questions" Noûs 1 1-8
- Chisholm, R (1973) "Parts as essential to their wholes" Review of Metaphysics 26 581–603
- Chisholm, R (1976) *Person and Object* Allen & Unwin
- Clayton, P and P Davies (2006a) *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion* Oxford University Press
- Clayton, Phillip (2006b) "Conceptual foundations of emergence theory" found in Clayton et al (2006)
- Clayton, P and S Kauffman (2006) "On Emergence, Agency, and Organization" Biology and Philosophy 21 4 501-521
- Davies, P (2010) *Information and the Nature of Reality: From Physics to Metaphysics* Cambridge University Press
- Davidson, D (1970) "Mental Events" found in Foster and Swanson (eds.) (1970) *Experience and Theory* Duckworth
- Dawkins, R (2005) *The Ancestors Tale* Houghton Mifflin
- Deacon, T (1997) *The Symbolic Species* W.W. Norton & Co.
- Debord, G (1967) *La société du spectacle* Buchet-Chastel
- DeLanda, M (1997) *A Thousand Years of Nonlinear History* Zone Books
- DeLanda, M (2002*) Intensive Science and Virtual Philosophy* Continuum Press
- Drummond, J (2003) "Husserl's third logical investigation: parts and wholes, founding connections, and the synthetic a priori"; found in *Husserl's Logical Investigations* Dahlstrom, J (ed) Springer]
- Dupré, J (1995) *The Disorder of Things: Metaphysical Foundations of the Disunity of Science* Harvard University Press

- Dupré, J (2008) *The Constituents of Life* (The Spinoza Lectures) Van Gorcum
- Earley, J (2005) “Why there is no salt in the sea” Foundations of Chemistry 7 1 85-102
- Ellis, G F R (2006) “On the nature of emergent reality” found in Clayton et al (2006) 79-109
- Fine, K (1995) “Part-whole” found in *The Cambridge Companion to Husserl* eds. B Smith and DW Smith, Cambridge University Press 463-485
- Foucault, M (1966) *The Order of Things: An Archaeology of the Human Sciences* Routledge
- Garber, D and J-B Rauzy (2004) “Leibniz on body, matter and extension” *Aristotelian Society* Supplementary Volume 78 1 23–40
- Gilson, E (1955) *History of Christian Philosophy in the Middle Ages* Random House
- Gould, SJ (2002) *The Structure of Evolutionary Theory* Harvard University Press
- Bennett, MR and PMS Hacker (2003) *Philosophical Foundations of Neuroscience* Wiley
- Bennett, MR and PMS Hacker (2008) *History of Cognitive Neuroscience* Wiley
- Hacking, I (1983) *Representing and Intervening, Introductory Topics in the Philosophy of Natural Science* Cambridge University Press
- Halliwell, JJ, J Pérez-Mercader, WH Zurek (eds) (1994) *Physical Origins of Time Asymmetry* Cambridge University Press
- Hamilton, E (1854) *The Works of Thomas Reid* Maclachlan & Stewart
- Harold, Franklin (2001) *The Way of the Cell: Molecules, Organisms, and the Order of Life* Oxford University Press
- Harman, G (2002) *Tool-Being: Heidegger and the Metaphysics of Objects* Open Court
- Harman, G (2009) *Prince of Networks: Bruno Latour and Metaphysics* Re.Press
- Harman, G (2010) *Towards Speculative Realism: Essays and Lectures* Zero Books
- Hartz, G (1992) “Leibniz's Phenomenalisms” *Philosophical Review* 101 3 511-549
- Haslanger, S (2003) “Persistence through time” in Loux, M and D Zimmerman (eds) (2003) *The Oxford Handbook of Metaphysics* Oxford University Press
- Hayek, F (1973) *Law, Legislation and Liberty* University of Chicago Press
- Heard D (2006) “A new problem for ontological emergence” The Philosophical Quarterly 56 222 55–62
- Heller, M (1984) "Temporal parts of four dimensional objects" Philosophical Studies 46 323-34
- Heller, M (1990) *The Ontology of Physical Objects* Cambridge University Press
- Heller, M (1992) “Things change” Philosophy and Phenomenological Research 52 3 695-704

- Hempel, C (1952) *Fundamentals of Concept Formation in Empirical Science* University of Chicago Press
- Hempel, C (1965) *Aspects of Scientific Explanation* Free Press
- Holland, J (2000) *Emergence: From Chaos to Order* Oxford University Press
- Hulswit, M (2005) “How causal is downward causation?” Journal for General Philosophy of Science 36 2 261-287
- Husserl, E (1970) *Logical Investigations* trans. JN Findlay; Routledge / London & Kegan Paul
- Janzen, G (2011) “On three arguments against endurantism” Metaphysica 12 2 101-115
- Kant, I (1910) *Monadologia Physica* in *Kant's Gesammelte Schriften*, Königlich Preussischen Akademie der Wissenschaften, I. Abt., Bd. I., Berlin, trans/ed G. Reimer 473-487
- Kauffman, S (1993) *Origins of Order* Oxford University Press
- Kauffman, S (1995) *At Home in the Universe* Oxford University Press
- Kauffman, S (2000) *Investigations* Oxford University Press
- Kauffman, S (2006) *Reinventing the Sacred* Basic Books
- Kauffman, S (2004) "Autonomous agents", found in Barrow, JD, P.C.W. Davies, and C.L. Harper Jr. (eds.) *Science and Ultimate Reality: Quantum Theory, Cosmology, and Complexity* Cambridge University Press
- Kemeny, J. and Oppenheim, P. (1956) “On reduction” Philosophical Studies, 7 6–17.
- Kim, J (1964) “Inference, Explanation and Prediction” The Journal of Philosophy 61 12 360-368
- Kim, J (1982) “Psychophysical Supervenience” Philosophical Studies 41 51-70
- Kim, J (1993a) *Supervenience and Mind* Cambridge University Press
- Kim, J (1993b) “Multiple realization and the metaphysics of reduction” in Supervenience and Mind ed. J. Kim (1992) Cambridge University Press
- Kim, J (1999a) “Making sense of emergence” Philosophical Studies 95 ½ 3-36
- Kim, J (1999b) *Mind in a Physical World* MIT Press,
- Kim, J (2005) *Physicalism, or Something Near Enough* Princeton University Press
- Kim, J (2006a) “Being realistic about emergence” found in Clayton et al (2006)
- Kim, J (2006b) “Emergence: core ideas and issues” Synthêse 151 547-559
- Kitcher, P S (1992) "Gene: Current Usages", in E. Keller and L Lloyd (eds.) *Keywords in Evolutionary Biology* Harvard University Press
- Klir, G and B Yuan (1995) *Fuzzy Sets and Fuzzy Logic: Theory and Applications* Prentice Hall
- Koslicki, K (2006) “Aristotle's mereology and the status of form” *Journal of Philosophy* 103 12 715-736
- Koslicki, K (2007) “Towards a Neo-Aristotelian mereology” *Dialectica* 61 1 127–159

- Krauss, L (2012) *A Universe from Nothing: Why There is Something Rather than Nothing* Free Press
- Kripke, S (1980) *Naming and Necessity* Blackwell
- Laughlin, RB (1999) "Fractional quantization" *Rev Mod Phys* 71 863
- Leibniz, G (1971) GM: "Mathematical Correspondence" or *Leibnizens Mathematische Schriften* ed. C. I. Gerhardt, 7 vols. (Berlin and Halle: Asher and Schmidt, 1849-63; reprint ed. Hildesheim: Olms)
- Leibniz, G (1961) GP: "Philosophical Correspondence" or *Die philosophischen Schriften von Leibniz*. trans/ed C. I. Gerhardt. Bd. 1-7. Berlin 1875-1890 (Neudruck: Hildesheim 1960-1961)
- Leibniz, G (1898) *Monadology* (1714) trans. R Latta
- Leibniz, G (1956) *Philosophical Papers and Letters* University of Chicago Press, (two volumes) trans/ed L Loemker
- Lehn, JM (1995) *Supramolecular Chemistry* VCH Verlag
- Lewes, GH (1875) *Problems of Life and Mind, Vol. 2* Kegan Paul, Trench, Turner, and Co
- Lewis, D (1986) *On The Plurality of Worlds* Blackwell
- Libardi, M (1994) "Applications and limits of mereology: from the theory of parts to the theory of wholes" *Axiomathes* 5 1 13-54
- Lucretius (1994) *De Rerum Natura* trans. William Ellery Leonard
- Luhmann, NJ (1993) "Deconstruction as Second-Order Observing" New Literary History 24 4 Papers from the Commonwealth Center for Literary and Cultural Change 763-782
- Maturana, H and F Varela (1980) *Autopoiesis and Cognition: The Realization of the Living* D. Reidel Publishing Company
- Maturana, H and F Varela (1987) *The Tree of Knowledge* New Science Library / Shambhala (rev. 1988, 1992)
- McGlaughlin, B (1992) "The rise and fall of British emergentism" in *Emergence or Reduction?* Beckermann, A, H Flohr, and J Kim, eds. Berlin:Walter de Gruyter 49-93
- McGlaughlin, B (1997) "Emergence and supervenience" Intellectica 1997/2 25 25-43
- McKinnon, N (2002) "The endurance/perdurance distinction" Australasian Journal of Philosophy 80 3 288-306
- Merlin, D (1991) *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition* Harvard University Press
- Merlin, D (2001) *A Mind So Rare: the Evolution of Human Consciousness* WW Norton
- Merricks, T (1999) "Persistence, parts, and presentism" Noûs 33 3 421-438
- Merricks, T (2001) *Objects and Persons* Oxford University Press
- Mill, JS (1843) *System of Logic* London: Longmans, Green, Reader, and Dyer (8th ed., 1872)
- Morgan, CL (1923) *Emergent Evolution* London: Williams and Norgate
- Morowitz, H (2002) *The Emergence of Everything* Oxford University Press

- Nagel, E (1952) “Wholes, Sums and Organic Unities” *Philosophical Studies* 3 2 17-32
- Nagel, E (1979) *The Structure of Science* Harcourt, Brace and World
- Penrose, R (1989) *The Emperor's New Mind* Oxford University Press
- Penrose, R (1994) *Shadows Of The Mind: A Search for the Missing Science of Consciousness* Vintage
- Pickering, A (1995) *The Mangle of Practice: Time, Agency and Science* University of Chicago Press
- Place, UT (1956) “Is Consciousness a Brain Process?” British Journal of Psychology 47 44–50
- Plato
- Plato (1999) *Parmenides* trans. B Jowett, The Softback Library [1871]
- Plato (2001) *Philebus* trans. B Jowett, Modern Library [1871]
- Plato (1999) *Sophist* trans. B Jowett, The Softback Library [1871]
- Plato (2001) *Thaetetus* trans. B Jowett, Modern Library [1871]
- Prigogine, I (1961) *Introduction to Thermodynamics of Irreversible Processes* Interscience
- Prigogine, I and G Nicolis (1977) *Self-Organization in Non-Equilibrium Systems* Wiley
- Prigogine, I and I Stengers (1984) *Order Out of Chaos* Bantam Doubleday Dell
- Prigogine, I (1997) *The End of Certainty* Free Press
- Protevi, J (2006) “Deleuze, Guattari, and emergence” Paragraph: A Journal of Modern Critical Theory 29 2 19-39
- Protevi, J (2009) “Inflections of Emergence and Politics in Francisco Varela” in Clarke, B and MBN Hansen (eds) *Emergence and Embodiment: New Essays on Second-Order Systems* Theory Duke University Press
- Rea, M (2003) “Four-dimensionalism” in Loux, M and D Zimmerman (eds) (2003) *The Oxford Handbook of Metaphysics* Oxford University Press
- Rescher, N and P Oppenheim (1955) “Logical analysis of gestalt concepts” British Journal for the Philosophy of Science, 6 22 89-106
- Sawyer, RK (2001) “Emergence in sociology: contemporary philosophy of mind and some implications for sociological theory” American Journal of Sociology 107 3 551-85
- Scott, A (1995) *Stairway to the Mind* Copernicus/Springer
- Searle, J (1984) *Minds, Brains, and Science* Harvard University Press
- Searle, J (1992) *The Rediscovery of the Mind* 1992 MIT Press
- Searle, J (2008) “Reductionism and the irreducibility of consciousness” in Bedau and Humphreys (2008)
- Sider, T (1993) "Van Inwagen and the Possibility of Gunk" Analysis 53 285–289
- Sider, T (1997) “Four-dimensionalism” Philosophical Review 106 197-231
- Sider, T (2001) *Four-Dimensionalism* Oxford University Press
- Silberstein, M (1998) “Emergence and the Mind-Body Problem,” Journal of Consciousness Studies, 5 4 464-482

- Silberstein, M and J McGeever (1999) “The search for ontological emergence” The Philosophical Quarterly, 49 195
- Simons, P (1987) *Parts: A Study in Ontology* Oxford University Press.
- Simons, P and J Melia (2000) “Continuants and occurrents” Proceedings of the Aristotelian Society, Supplementary Volumes 74 59-75 + 77-92
- Simons, P and W Baumgartner (1994) “Brentano's mereology” *Axiomathes* 5 1 55-76
- Smart, JJC (1959) “Sensations and Brain Processes” Philosophical Review 68 141–156
- Smith, B (1998) “The basic tools of formal ontology” found in Nicola Guarino, ed. (1998) *Formal Ontology in Information Systems* IOS Press (Frontiers in Artificial Intelligence and Applications), Washington, D.C. 19–28
- Smith, B and K Mulligan (1982) "Pieces of a Theory" found in Smith, B (ed) *Parts and Moments: Studies in Logic and Formal Ontology*, Philosophia, Munchen 15-109
- Sperry, R (1964) “Problems outstanding in the evolution of brain function” *James Arthur Lecture* American Museum of Natural History
- Sperry, R (1969) “A modified concept of consciousness” Psychological Review 76 532-536
- Sperry, R (1978) “Mental phenomena as causal determinants in brain function” in Globus, P (1978) *Consciousness and the Brain* Plenum
- Sperry, R (1986) “Macro- versus micro-determinism” *Philosophy of Science* 53 2 265-270
- Thomas Aquinas (1954) *The Disputed Questions on Truth* (Questions XXI – XXIX) trans. RW Schmidt, Henry Regner Company
- Thomas Aquinas (1968) *On Being and Essence* trans. (and notes) Armand Maurer, Pontifical Institute of Mediaeval Studies Toronto
- Thomas Aquinas (1981) *Summa Theologiae* trans. by English Dominicans (London: Burns, Oates, and Washbourne 1912-36; repr. New York: Benziger 1947-48; repr. Christian Classics
- Thomas Aquinas (1997) *Commentary on The Sentences* (excerpt) trans. R McInerny, in *The Basic Writings of Thomas Aquinas* Hackett Publishing
- Thomson, JJ (1983) “Parthood and identity across time” The Journal of Philosophy 80 201-220
- Tymieniecka, AT (2002) *Phenomenology world-wide: foundations, expanding dynamisms, life-engagements: a guide for research and study* Springer
- Van Gulick, R (1985) “Physicalism and the subjectivity of the mental” Philosophical Topics 12 51–70.
- Van Gulick, R (1993) “Understanding the phenomenal mind. Are we all just armadillos?” in *Consciousness*, ed. M. Davies and G. Humphreys, Blackwell
- Van Gulick, R (1999) “Conceiving beyond our means: the limits of thought experiments” in *Toward a Science of Consciousness III*, ed. S. Hameroff, A. Kazniak and D. Chalmers, MIT Press

- Van Gulick, R (2001) “Reduction, emergence and other recent options on the mind/body problem” Journal of Consciousness Studies, 8 9-10 1-34
- Varzi, A (1996) “Parts, wholes and part-whole relations: the prospects of mereotopology” *Data and Knowledge Engineering* 20 259–286
- Wasserman, R (2006) “The problem of change” Philosophy Compass 1 1 48-57
- Weinberg, S (1979) *The First Three Minutes: A Modern View of the Origin of the Universe* Basic Books (rev. 1993)
- Weinberg, S (1993) *Dreams Of A Final Theory: The Search for The Fundamental Laws of Nature: Search for the Ultimate Laws of Nature* Vintage
- Whitehead, AN (1920) *The Concept of Nature* CreateSpace Independent Publishing [2012]
- Whitehead, AN (1925) *Science and the Modern World* Cambridge University Press
- Whitehead, AN (1929) *Process and Reality* Macmillan
- Whitehead, AN (1933) *Adventures of Ideas* Macmillan
- Wiener, N (1948) *Cybernetics: Or Control and Communication in the Animal and the Machine* MIT Press (rev.1961)
- Wiener, N (1958) *Nonlinear Problems in Random Theory* MIT Press
- Wiggins, D (1980) *Sameness and Substance* Harvard University Press
- Williams, JRG (2006) "Illusions of gunk" Philosophical Perspectives 20 1 493–513
- Williams, JRG (2007) “The Possibility of Onion Worlds: Rebutting an argument for structural universals” The Australasian Journal of Philosophy, 85 2 193-203
- Wolfram, S (2002) *A New Kind of Science* Wolfram Media
- Zeh, H D (1989) *The physical basis of the direction of time* Springer