Synthese (2009) 166:453–477
DOI 10.1007/s11229-008-9372-0

A process-based model for an interactive ontology

Richard Campbell

Received: 16 December 2007 / Accepted: 11 July 2008 / Published online: 27 August 2008

Abstract The paper proposes a process-based model for an ontology that encompasses the emergence of process systems generated by increasingly complex levels of organization. Starting with a division of processes into those that are persistent and those that are fleeting, the model builds through a series of exclusive and exhaustive disjunctions. The crucial distinction is between those persistent and cohesive systems that are energy wells, and those that are far-from-equilibrium. The latter are necessarily open; they can persist only by interaction with their environments. Further distinctions, developed by means of the notions of self-maintenance and error detection, lead to the identification of complex biological organisms that are flexible learners, some of which are self-conscious and form themselves into social institutions. This model provides a non-reductive model for understanding human beings as both embodied and yet emergent. In particular, it provides a way of characterizing action as 'metaphysically deep', not an ontological embarrassment within an otherwise physicalist world.

Keywords Ontology · Process metaphysics · Physicalism · Emergence · Self-maintenance

How are we to understand ourselves and the world? Traditionally, philosophers have proposed ontological schemes as if they were sitting on Mount Olympus, looking on with divine detachment at the world they were describing. Their conceit was to pretend that they were not themselves part of that same world. They talk as if they are not rooted in the flow of its history, nor subject to the limitations of any specific context. But that conceit is no longer plausible; we all now know ourselves to speak

R. Campbell (✉)
Department of Philosophy, The Australian National University, Canberra 0200, Australia
e-mail: rcampbell@netspeed.com.au; Richard.Campbell@anu.edu.au

from some specific historical and cultural context. Accordingly, it is more honest, and more credible, to understand any ontological scheme that we might propose as simply a *model*, a way of understanding how the processes and items in the world have come to be structured. Despite the Olympian pretensions of earlier (and some contemporary!) philosophers, their ontological schemes bespeak their own times and places in cultural history. Even if what we propose today is the most comprehensive and the most scientifically well-informed that we can muster, it still will be at best just a *model*, likely to be superseded in future generations by more adequate models. So why not acknowledge that? What follows is proposed in that spirit.

I present here a metaphysical model as a radical alternative to the reductive micro-physicalist picture generally assumed by popular writers and philosophers alike.[1] Their vision is that "all the things that exist are physical things—either basic bits of matter or made up of bits of matter" (Kim 2002, p. 640).[2] Or as Barry Loewer puts it, "Physicalism claims that all facts obtain *in virtue of* the distribution of the fundamental entities and properties—whatever they turn out to be—of completed *fundamental physics*" (Loewer 2001, p. 37). All that happens in the world is determined ultimately by the behaviour of these 'basic bits of matter' that are the fundamental entities of the world. These micro-physical entities are clearly some sort of particular. But what sort? Some physicalists believe that it does not matter, that the notion of a physical particular might be defined as an object, a concrete event, or whatever (Jackson 1998, p. 6). Howbeit, the key commitment of physicalism is to some kind of *basic particulars*, which are the fundamental constituents out of which everything in the world is composed, and whose properties and relations are sufficient to determine everything that is true about the world.[3]

This doctrine has its roots in John Locke, who began his discussion of General Terms in his *An Essay Concerning Human Understanding* with a brazen assertion: "All things, that exist, being Particulars, . . ." (Locke 1975, III, iii, 1). To posit such a starting-point, without the slightest argument, was breath-taking in its presumption. After all, the ontological status of universals, and the relation of particulars to them, had been one of the most hotly contested philosophical issues in the preceding centuries. Nevertheless, for Locke, Isaac Newton, Robert Boyle, and their Seventeenth century contemporaries, those medieval debates were simply swept aside. All that exists, they blithely assumed, are particulars: 'corpuscles' and aggregates of them, solid lumps of matter. That ontological model, as easy to visualize and comprehend as a game of billiards, still holds the imagination of many present-day philosophers in its thrall. Despite

---

[1] The ideas assembled in this paper have been gleaned from many sources, but my main debts are to Mark Bickhard, Johanna Seibt, Cliff Hooker and Wayne Christensen.

[2] The same thought is repeated in his more recent *Physicalism, or Something Near Enough* (Kim 2005, p. 7).

[3] Actually, the concept of 'basic particular', is more problematic than this statement discloses. The concept of a 'particular' is clearest as a term derived from the logic of discourse, that is, the particulars are what the quantifiers ("every" and "some") in some discourse range over. In that sense, 'basic particulars' would be those items which *that discourse* treats as not logically derivable from any others. In that sense, psychology would have its own 'basic particulars'. Obviously, that is not strong enough for a physicalist! Yet if the latter concedes that, for example, mental items cannot be *logically derived* from descriptions of physical items, the latter would have to be 'basic particulars' in some other sense, yet to be clarified.

the undermining of Newtonian physics a century ago, both by relativity theory and by quantum theory, that model persists in the continuing talk of *elementary particles*. This is despite the fact that the word "particle" can only be used in the loosest possible way when one is describing quantum-level phenomena, for these phenomena can no longer be thought of as tiny corpuscles. What lies behind the persistence of this talk seems to be an inability to conceive of how the middle-sized things that furnish our everyday world could be composed of anything other than smaller things, micro billiard balls.

1 Processes

But there is a problem: contemporary physics tells us that *there are no elementary particles*. Quantum field theory shifts the basic ontology of the universe from micro-particles to quantum fields. If nothing is fundamentally corpuscular, that suggests that all things have to be conceived fundamentally as *processes* of various scales and complexity, having causal efficacy in themselves. What have seemed to be particles are now reconceptualized as particle-like processes and interactions resulting from the quantization of field processes and interactions, and those are no more particles than are the integer number of oscillatory waves produced by plucking a guitar string. Each of these so-called things is a quantized field process. The world consists of *organized fields* in process—all the way down, and all the way up (Bickhard 1998b, 2000).

Should that worry physicalists? Why could they not say: Well, the world might be composed of processes, but if so, all that shows is that the basic particulars are particular processes occurring at some place and time?

That response, however, will not do. There have been a few philosophers who have tried to explore the ontology of processes, but even fewer have been able to escape the prejudice in favour of particulars. Whenever processes have been taken seriously, attention has been focussed upon localized processes: the specific performance of a human activity, such as my reading the newspaper this morning, or the occurrence of some process in a specified spatiotemporal region, such as last Sunday's rain in Canberra. But to take localized processes as providing the model for basic particulars is to misunderstand the logic of processes (Seibt 1996, 1997, 2001, 2003).[4] Let me briefly explain.

To bring the character of processes to light, it is helpful to compare them to kinds of stuff, out of which different things can be made. Just as it is right to recognize kinds of stuff, designated by mass terms, so we have to recognize generic processes—such as running, raining, fire and light—which can occur in different places and times. Once we free our imaginations from the prejudices of metaphysical orthodoxies, we can easily see that there is nothing conceptually peculiar about the concept of generic processes. Running, raining, fire and light are familiar processes. They can occur in different regions, and can re-occur. That is, while rain must occur in some spatiotemporal region in order to exist at all, raining can occur both in Canberra and in Cambridge,

---

[4] Johanna Seibt has since 1990 been developing the most carefully worked out account of the logic of processes, in what she initially called Free Process Theory and now calls General Process Theory. See her contribution to this issue.

both yesterday and on Tuesday last week, whereas yesterday's raining in Canberra, which is a specific process, cannot occur in Cambridge and cannot have occurred last week. Generic processes are dynamic features, a 'going on thusly', something that is not a particular in the traditional sense at all (Seibt 2001). Different kinds of process are readily distinguished—snowing is clearly not the same as raining—but those distinctions do not make reference to specific spatiotemporal regions (which is why I call them 'generic').

All right, but what about localized processes? Are they not particulars? The most helpful way to see why not is to consider the logic of process-descriptions. The fundamental way of describing processes is by those words standardly classified as 'activity verbs', which often take a 'continuous' form, such as "read" in "she is reading a novel", and "rain" in "it was raining all morning". Indeed, developing a process-based ontology is a way of according proper primacy to verbs, in contrast to the traditional priority that philosophers have given to those nouns which designate substantial things.

Nevertheless, it is a mistake to assume that the ontological category implied by an expression is determined by its *grammatical* category. Some nouns, such as "rain" and "light" refer to processes, not to things. Nor are those category implications wholly determined by the *lexical* meaning of the words used. Although verbs are usually classified lexically as expressing a certain 'default' occurrence type ("run" expresses an activity, "win" an event), their crucial feature is what linguists call their 'aspect'. The context of use can shift the aspectual meaning of verbs from their 'default' occurrence type to express happenings of a different type. Thus, while running is a process, in "John ran three races" the verb "ran" has shifted aspect in order to denote (three) events. That said, it will be convenient to call 'activity verbs' those verbs whose aspect, either by default or in context, is to express an activity.

A standard transformation on activity verbs yields verbal nouns, or gerunds, such as "reading" in "the art of reading", and "running" and "walking" in "running is quicker than walking". This enables one of the most common ways of referring to processes. And certain nouns, such as "light", "rain", and "fire" also refer to processes (As I am using these words, a 'process' is referred to by such nouns, whereas the basic way of describing an 'activity' is by use of an activity verb. Of course, the word "activity" is itself a noun referring to some process involving the exertion of energy).

Now, there are significant logical similarities between verbs expressing processes and those nouns that denote kinds of stuff: mass nouns. A sentence in which some process is predicated can give rise to a nominalization that requires mass-quantifiers, not the quantifiers over denumerable domains of standard predicate logic (Taylor 1997; Mourelatos 1978). Hence, they are typically qualified by distinctive quantitative adverbial phrases. Consider "Jack used wood to build his house". A passive transformation of that sentence yields "Wood was used by Jack to build his house". We can now ask "How much wood was used by Jack?". Logically possible answers are: "a lot"; "a little"; "100 kilograms", etc. But it makes no sense to ask "How many?". On the other hand, quantifiers over things—particulars—presuppose that the relevant question is *how many* of the things referred to by the subject-term the predicate applies to: some or every one. Likewise, certain adverbial phrases, introducing measures, can be added to clauses describing processes to form new complete sentences. Examples of such quantitative adverbial phrases are "much", "a little", "a lot", and "for several hours", etc. (Roeper 1987). Accordingly, it makes logical sense to add "for several hours" to "it was raining", thus yielding "It was raining for several hours". And if Ken has been swimming we can ask "For how long did Ken's swimming last". It makes no logical sense to ask "How many rain?" or "How many was Ken swimming" (although if Ken was swimming in an Olympic-style pool, we can ask "*How many laps* did Ken swim?", but then we have changed the aspect of "swim" and are asking a different question). This difference is one crucial respect in which the logic of processes differs fundamentally from the logic of things.

Another way that adverbial phrases can be added to clauses describing generic processes is to locate them in specific periods of times and at specific places. For this reason, there is a straightforward inference from a sentence asserting the occurrence of a specific process in a time and place to a sentence asserting the occurrence of the relevant generic process. Thus, for it to happen that Ken was swimming in the sea at Malua Bay this morning is for it to happen that he was swimming this morning, that he was swimming in the sea, that he was swimming, etc.

The fact that 'adverb dropping' like this yields valid inferences clearly requires logical explanation. Unfortunately, excessive preoccupation with particulars has led philosophers to attempt to assimilate events and processes to particulars, even though no-one would dream that events and processes are anything like Aristotelian substances. One suggestion that some have found attractive is to propose that "Ken is swimming" should be analysed as "There is a swimming by Ken", which lends itself easily to being assimilated to first-order predicate logic with identity. Then "Ken was swimming in the sea at Malua Bay" would be analysed as "There was a swimming by Ken, and it happened in the sea at Malua Bay", which again can be easily assimilated to first-order predicate logic and could accordingly support the inferences noted in the previous paragraph (as was proposed by Davidson, 1980). But not only is this analysis very forced, it is seriously misleading. In particular, "Ken is swimming" and "There is a swimming by Ken" have different ontological implications. The second contains a reference to certain processes, whereas "Ken is swimming" does not *refer* to any process. The two sentences are indeed logically equivalent, in that they have the same truth-conditions, but they are *not* semantically identical. A similar situation obtains with the pair of sentences "Ken is a swimmer" and "Ken belongs to the class of swimmers" and with the pair "There are more As than Bs" and "The number of As is greater than the number of Bs". The second in each of these pairs refer to classes and numbers, respectively, whereas the first sentence in each pair does not.

Yet another crucial difference between things and generic processes is the way the part-whole relationship applies to them. A part of a table is not a table, nor is a part of a dog a dog. In contrast, like the kinds of stuff designated by mass terms, the parts of a process—whether generic or specific—are homomerous, or 'like-parted'. That is, just as (almost) all parts of water is water, so (almost) all parts of raining is raining. Now, despite these differences, in traditional ontologies the mereological (part-whole) relation is taken to be transitive. That is, although the parts of a table are not themselves tables, a table can be decomposed into its parts (say, four legs and a top), and those parts can be decomposed into yet smaller parts. By transitivity, those smaller parts are also parts of the table. The transitivity of the part-whole relation, which is a distinctive principle of classical extensional mereology, is crucial to physicalism. Entities belonging to a given level, except those at the very bottom, are taken to have an exhaustive decomposition, without remainder, into entities belonging at a lower level, and so on until one reaches the alleged basic physical particulars (Entities at the very bottom are supposed not to have physically significant proper parts.). As a consequence of this 'mereological collapse', physicalism supposes that it is these basic particulars that determine all the properties and powers of the entities which they compose.

But activities and stuffs have functional roles that disrupt this sort of transitivity. One activity can be part of—that is, it is involved in, belongs to, or comes with—another activity, and the latter can be part of a third, but it does not follow that the first is necessarily part of the third.[5] The transitivity principle of classical mereology—'is *a* part of'—can be defined in terms of the more generic relation 'is part of', but not vice versa (Seibt 2001).

Focussing on localized processes does not obliterate these differences in logic. Specific processes are not denumerable in a non-arbitrary way. They can merge and separate into branches. Consider the example of fire, which, as Heraclitus first realized, is the very paradigm of a process. In January 2003, lightning strikes caused bushfires to break out in four distinct places in the mountains to the west of Canberra. On 18 January, strong north-west winds whipped them up so that they merged into a firestorm which engulfed the western suburbs of the city, destroying nearly 500 homes. Was the subsequent judicial inquiry dealing with four fires, or just one? To press the "how many?" question would simply generate conceptual confusion. Whilst any specific process begins at some time, and ends sometime later, many are not sufficiently hard-edged and well-defined to count as any sort of particular. So, while generic processes only ever occur as localized in periods of time and regions of space, and while those specific processes can be referred to and quantified over in a logic analogous to that of mass terms, that very logic demonstrates that no sub-class of them is suitable to be considered as basic particulars.

A further misunderstanding has been the tendency to think of specific processes as composed of a series of particular events, and thereby to propose an atomistic reduction of them. But there is now no reason, other than prejudice, to do so, and good reasons to reject the assumption that processes are constituted by a series of events.[6] Processes take time, and like time itself, are generally continuous. That is not to deny that particular events might occur at various stages in some complex process, but the process itself cannot be thought of as a concatenation of such events. To think so is to make the same mistake as supposing that the continuum can be constructed out of a concatenation of points.

There is a similar manoeuvre that might tempt some physicalists. Suppose they concede that specific processes are not apt to be considered as particulars, but shift

---

[5] Seibt (2001) gives the following example (Footnote 2): Changing diapers is part of being a parent. Opening the box with wipes is part of changing diapers. Pressing the thumb upwards is part of opening the box with wipes. But it does not follow that pressing the thumb upwards is part of being a parent. For her analysis of the ontology of processes, see also Seibt (2003) and her paper in this issue.

[6] In fairness, we should acknowledge that when A.N. Whitehead attempted in the 1920s to articulate a process metaphysics generated out of atomic events, he believed that the discrete character of the new quantum mechanics of his day did require processes to have an atomic constitution.

their ground to the claim that specific processes fulfil the role they once envisaged for basic particulars. That is, they now might claim that certain specific processes are the fundamental constituents out of which everything in the world is composed, and that their properties and relations are sufficient to determine everything that is true about the world. But this manoeuvre will not rescue their position. For if the world consists of organized fields in process, all the way down and all the way up, it follows that *no processes are basic* in the physicalists' sense, and therefore there is no privileged class of fundamental processes upon which their physicalism could be based.

But why should we think that the world consists of processes all the way down and all the way up? Simply because space-time is continuous! Since contemporary physics requires that quantum fields be taken seriously, and since the only coherent way to conceptualize quantum fields is as processes extended in space-time, it follows that *any* process, no matter how micro, consists of yet smaller processes, ad infinitum.[7] This argument, grounded in contemporary physics, is independent of the previous arguments above. It therefore strengthens the case for rejecting the physicalists' prejudice in favour of 'basic particulars'.

Physicalists might have one last try. They might accept that there are no base-level processes, but try to rescue physicalism along the following lines. Suppose we restrict our view just to those processes discernible at the finest level of resolution, however small that level might be. Why cannot physicalists claim, in the light of contemporary physics, that these fine-grained processes are what fulfil the role previously claimed for basic particulars? The physicalist issue simply is: Are all the truths determined at that level sufficient, in principle, to determine all the truths in the world? If so, that is all physicalists need.

This riposte is deeply ambiguous; the answer to the question depends upon whether those truths determined at the level of finest resolution include all the *relational* truths. The physicalist is committed to this being so. But not all the relevant relational truths are discernible at the 'finest level of resolution'; some, probably most, of the patterns that are causally significant are of a larger scale than that. If we are restricted to just those that are of a small enough scale to be *discernible* at such a fine grain, then the answer to the question must be negative. Reference to 'all the truths determined at the finest level of resolution' glosses over this crucial issue of the *scale* of causally efficacious patterns. If these truths are not restricted to patterns of a small enough scale

---

[7] David Lewis (1986a,b), expounding 'Humean supervenience', seems to think that fields could be defined at each point (presumably in space-time), and that these points, or the point-sized occupants of such points, are themselves particulars. If that were possible, it would get around the argument above, but his proposal is simply false. Mark Bickhard has pointed out that Lewis' proposal is only correct if those points are in a topology (metric, etc.), but that topology cannot yield a field if those points are individuated. In a point set topology, most or all of the points are defined relatively, not absolutely. A field cannot be defined point by point. Continuous topological 'properties' are ***relational*** 'properties', and the continuum requires a continuum of points, each of which has only an implicit, relational ontology. It *is* possible to provide an equation that gives a field density point by point, but that can only be done by borrowing the metric and topological properties of the underlying space-time. And space-time is not 'Humean'. It might be that space-time can be construed in terms of points (Einstein did), but they are not Humean points. They are not particulars; they don't have independent existence; they cannot all be named; they cannot all be identified, etc. Lewis' favourite analogy of a newspaper photograph, in which a picture is determined by the distribution of tiny dots, is seriously misleading. Those dots, no matter how tiny, are not point-like; each has extension.

to be discernible at such a fine grain, we may include relational truths at the scale of a whole human person—or larger. Then the answer to our question is clearly yes, but the position cannot claim any longer to be a serious version of *physicalism*.

Generic processes are what scientific theories typically describe. Biologists have been clear about this for many years, and as physicists shake themselves free of Newtonian models, they are coming to see it too. Physicalist philosophers are just out-of-date. What is physically significant about processes is how they are configured and organized; their configurations and organizations are what ground their causal powers, not the properties borne by particular things which might be supposed to constitute them.

Accordingly, this paper takes processes as the basic ontological category, and upon that basis builds up, step by step, a model of ontological kinds, through a series of disjunctions. Articulating the model in this way might give the appearance of being *a priori*, but that is only superficial. The selection of these disjunctions is informed by the outcomes of many empirical investigations. Unlike the a priori dogmatism of physicalists, this model is empirically grounded.

2 Persistence

Processes exist *only* in some organization or other. There is nothing to a field without its organization. And it is their organization which generates their properties, and does all their causal work. While any system of processes organizes its sub-processes into some dynamic pattern, those constituents are *altered* by their coming to play a role in the larger whole.

Now, some organizations of process are fleeting, such as Newton's legendary falling apple. Others are persistent, or at least reasonably persistent—indeed, some endure for eons. By "persistent" here I mean nothing more than that they endure through significant changes in their environments. In most cases, of course, there is an intrinsic reason why an organization of processes persist through environmental changes—for example, their organization might prove to be cohesive. But such intrinsic properties of an organization is not being invoked at this stage of development of the model. All that is being invoked is persistence through environmental change. In this sense, persistence is a relative quality; it turns on the organized process lasting for a longer time-span than the other process in its surroundings. These simple observations yield our first crucial disjunction: *either the organization of a process is persistent relative to changes in its environment, or it is not.* That distinction is the first step in this model.

3 Cohesion

Some persistent processes are stable in certain ways, but not in others. In particular, some groups of processes manifest stability in ways which are sufficient to demarcate them from their environment as integral systems, whereas others exhibit certain sorts of stability, but do not constitute cohesive systems. Consider a group of gas molecules: it assumes whatever shape and condition its containing environment imposes, and it will simply disperse if it is not constrained by a closed container. Whilst most gases are chemically stable, the group of gas molecules does not manifest any overall integration; a quantity of a gas has no internal cohesion. But there are many kinds of processes that do, and which consequently endure not only over time but also against considerable perturbation. Accordingly, the next step is to distinguish, amongst persistent processes that manifest some sort of stability, *those that constitute cohesive systems from those that do not.*

A *cohesive* system is one in which its various internal processes work together to ensure that one of the forms of stability which it manifests is spatio-temporal integrity (unlike a quantity of gas). The most significant kinds of cohesive systems are those in which the different elements are held together by dynamical bonds between them, which have the effect of individuating the system from its environment.[8] Accordingly, wherever we find such a system, we are able to identify and re-identify it. That is what licenses our calling it an entity.[9] And because certain process systems do cohere so as to constitute entities, we can count them; the question "how many?" makes sense.

That entities are not basic, but constituted, marks the crucial difference between our metaphysical model and that which has dominated Western philosophy since the time of the ancient Greeks. In the latter, as Aristotle systematized it, the basic ontological category is that of entity (*ousia*)—or, as modern English has modified the medieval Latin word used to translate *ousia*, substance. Contemporary physicalism, by still trying to promote 'basic particulars', is simply perpetuating a remnant of this metaphysical tradition; they are Aristotelian substances writ very small, with most of Aristotle's conceptual machinery stripped off. By contrast, in our model *entities* are not basic; they are derived. The basic category is that of process, and entities are certain kinds of persistent, cohesive processes.

What makes component processes into a strongly cohesive system—into an identifiable entity—are the internal bonds which constrain the behaviour of its constituent sub-processes in such a way that the totality behaves dynamically as an integral whole. These bonds arise from, although they are not reducible to, features of those quantum processes that constitute all of the components of a system. For example, the molecular bonds in the crystal lattice of a rock cause the rock as a whole to behave as a unified system under a large range of interactions; if it is kicked with moderate force, it moves relative to the ground. Contrast what happens when a pile of sand is kicked; the causal interactions between the grains of sand do not form bonds strong enough for the pile to behave as an integral system when it is kicked. Consequently, it scatters.

Of course, any system coheres only within a limited range of conditions. Hit the rock hard enough with a hammer and it will fracture; its cohesion will be disrupted.

---

[8] The notion that systems can be individuated in a principled way has been dismissed as ad hoc, arbitrary and observer-relative. But there are many causal properties that can serve as a basis for principled specifications of system identity that are not observer-dependent (Collier 1998; Christensen and Bickhard 2002).

[9] In general, we most readily identity and re-identify as entities those process systems whose internal bonds are strongly cohesive. These are the paradigm cases. But, as always, there are borderline cases. I later discuss a candle flame, which is marginally cohesive in the sense that the interaction of its internal processes enables a certain spatio-temporal integrity. But the denumerability of candle flames stems from the denumerability of the candles which feed them. In general, as discussed briefly above, fires are not denumerable entities in any strong sense.

The cohesion conditions of any individual entity can be specified physically (for example, a rock is cohesive within a specifiable range of temperature and external forces). Whether a given system is cohesive with respect to the forces to which it is subject is a determinate matter.

Now, the property of cohesion generates further causally significant properties. The internal bonds that cause the particles in a rock to hold together generate its mass, which is a holistic property. Likewise, it is the cohesion of a kite that allows it to fly, because the integrity of the kite's structure acts to sum the forces of the small interactions of air particles against it, generating a net lift force.

However cohesive systems (i.e., entities) are formed, they typically manifest properties that are different from those of their internal sub-processes. Some of these new properties result from an aggregation of the properties of the processes that are their constituents. For instance, the mass of a table is the aggregate of the mass of its four legs plus the mass of its top; they in turn are the aggregates of the mass of their constituent molecules. In such cases, the properties of the macro-level combination can be explained by an exhaustive and exclusive decomposition of the system into its proper parts. Nevertheless, it is important to note that more is involved in being a cohesive and causally effective aggregate than simply the arithmetical sum (that is, a bare conjunction) of its constituents. The components have to *stick together*, somehow or other, in order to effect a difference. And sticking together requires internal bonds.

Now, although certain properties of cohesive systems can be explained as resulting from aggregations of their internal parts, some systems are such that some of their properties cannot be so explained. There are at least four different conditions under which a system property may be an aggregate of the properties of its proper parts. I do not have space to elaborate them here. But the failure of system properties to satisfy one or more of these conditions provides precise and distinct senses in which they can be said to be 'more than the sum of the parts' of that system (Wimsatt 1986). It is evident that there are many macro-entities whose properties cannot be understood at all in terms of aggregation. These non-aggregative properties are crucial to causal emergence. Contemporary science now understands a good deal more about how different kinds of bonds organize their constituent processes into systems of significantly different kinds. Not all cohesive systems result from static bonding as rocks do; others, such as living cells, involve more dynamical relationships.

Yet even in the case of rocks, many of whose properties *can* be explained by aggregating the properties of the molecules they are made of, those micro-components themselves exhibit properties that cannot in turn be derived by aggregation of any kind from their internal sub-processes; they are emergent. Yet those components are also persistent, cohesive systems. A molecule of silica, for example, also exhibits cohesion, but it has emergent properties—unlike the properties of a rock itself.

The critical difference is between those systems whose cohesion is produced by bonds that have aggregative effects and those whose cohesion is produced by dynamical bonds that have non-aggregative, non-linear effects. Combinations of the latter kind bring into being new quantum field organizations with novel properties. The key point is that the fusion involved produces new unified wholes, with causal powers that cannot be derived by simply referring to the separate causal powers of its constituents, considered apart. It is the role of the empirical sciences to explicate more precisely the ways in which these bonding processes produce unified entities with novel properties, but the general phenomenon they seek to explicate is not hostage to any specific scientific theory. What is ontologically significant is that, in these cases of non-linear unification, the properties of the whole are somehow ‘more’ than the arithmetical sum of its parts—such system properties, and the causal powers of such a system, are *emergent*. Emergence should no longer be viewed as a dubious metaphysical mystery, but as explicable in terms of non-linear functions.

It follows that there is a simple link between cohesion and emergence: whenever a complex of processes organizes itself into a new cohesive system by forming internal bonds that involve non-linear forces, the resultant entity has emergent properties and powers. The result is the familiar picture of a multi-layered model of the world as stratified into different levels, in a micro-to-macro hierarchy. The Cartesian model of two substances—mind and matter—has long been outdated, but the usual response is to reject just one (usually mind). Thereby materialism, or physicalism as this philosophical position is more often called these days, simply perpetuates the Cartesian framework. I call it a ‘one-legged’ version of Cartesianism. We need a new model of Nature which genuinely moves beyond Cartesianism altogether. In this new model of Nature, entities, characterized by their distinctive properties and processes, *emerge* out of the processes which constitute the entities, properties, and processes of the levels below it. At the bottom is a level consisting of quantum fields, or whatever our best physics in the future tells us are the basic constituents out of which our world is generated. As we go up the levels, we successively encounter atoms, molecules, cells, multi-cellular organisms, human beings, social groups and institutions, etc.

4 Energy-well and far-from-equilibrium stability

The next step focuses on those cohesive systems whose dynamical bonds generate non-linear (that is, non-aggregative, emergent) properties. In many cases—for instance, all biological organisms—the integrity and cohesion of an entity depends not only on the dynamical bonds between its constituent elements but also on other internal and external operations. The processes that constitute cohesive, relatively persistent systems do not all operate in the same way; in fact, the resultant entities form a range that can be characterized by the two fundamental types that provide its endpoints. These two types of entity manifest ontologically different forms of stability. They are: (1) energy well stability; and (2) far-from-equilibrium stability.

‘Energy wells’ are cohesive process systems which persist at or near thermodynamic equilibrium, and whose organization can be disrupted only by an input, from external sources, of a critical level of energy. Typically, such a disruption of their organizational structure can only be brought about by a higher level of energy than they typically encounter in their ambient environment. Hence they are very persistent, cohesive, and robust.

Atoms are straightforward examples; they are a furious process of electron waves around an even more furious dance of quarks and gluons. In general, atomic processes are strongly cohesive and can be destabilized only by being bombarded by a great deal of external energy.

Combinations of such stable ‘energy well’ processes exist at the macroscopic level, yielding new, larger entities. For example, where different kinds of atoms interact in certain stable ways, they produce molecules with significantly different properties from those of the kinds of atoms that are their constituents. The organizational structure of hydrogen and oxygen atoms are such that two electron waves belonging to hydrogen atoms can come to participate in the outer ‘shell’ of an oxygen atom to produce a molecule of a new stable compound, H_2O. The resulting molecule of water has very different properties from the kinds of atoms that are its constituents. Strictly speaking, the constituent atoms of hydrogen and oxygen no longer exist. The configuration of their quantum fields has actually *changed* as a result of this fusion of their respective configurations, considered separately. They have been *transformed* into a new field, with quite new properties.[10]

The distinctive and most original feature of the model being articulated here, however, is its extended account of the *second* kind of cohesive and stable organization with emergent properties. The phenomenon of *far-from-equilibrium stability* poses a significant puzzle: how is it possible? How could an organized process that is not in thermodynamic equilibrium not only persist for some significant period without moving to equilibrium, but also exhibit a robust form of cohesion in the face of environmental changes? That is the ontologically revealing question to ask.

The most primitive kind of stable process of this kind is a system that is kept going by external means. An obvious example of a system that is maintained in a state of far-from-equilibrium stability largely in this way is planet Earth itself. Since far-from-equilibrium stability manifestly exists, its maintenance has to be a function of its being located within an interactive system of some sort. In the case of the earth, this is primarily a matter of energy flow from the sun to the earth and heat radiated from the earth into space.

Another example is a chemical bath in which interesting processes can persist because external pumps maintain a flow of the required chemicals from external reservoirs into the bath, while other pumps remove waste products. Until such a system is switched off, or runs out of chemicals, the chemical processes within the bath are sustained, but their persistence is completely dependent upon its environmental conditions: the pumps and the supplies contained in the external reservoirs. Such a chemical bath is, of course, a human artefact. As we will shortly see, there are more significant instances of far-from-equilibrium stability which occur naturally. But what enables any far-from-equilibrium system, whether natural or artificial, to survive are the ways its intrinsic processes keep interacting with its ambient environment. Its very persistence depends upon external supplies. In short, the stability of far-from-equilibrium processes is a function of their being *necessarily open* processes.

---

[10] For a more detailed discussion of such fusion, see Humphreys (1996). He argues that this fusion is to be understood in terms of the *replacement* of property instances. Phase changes such as this are still not well understood. If the transitions are through a critical point then it seems that it cannot be computed by means of dynamical equations because fluctuations occur on every scale simultaneously (I am indebted to Cliff Hooker for this comment on Humphreys use of “replacement”).

5 Self-maintenant systems

A chemical bath like that just described contributes nothing to the persistence of the conditions upon which it depends for its own continuance. But there are many other relatively stable far-from-equilibrium systems that do. Here is the next exclusive and exhaustive disjunction invoked by the model: *such systems either contribute to the persistence of the conditions upon which they depend, or they do not*. In terms of the development of the model, a candle flame is a familiar example that exhibits this distinctive difference from our simple chemical bath. For a candle flame is a complex of processes that make several active contributions to its own persistence, including its maintaining a (sometimes flickering) spatio-temporal integrity. Most importantly, a candle flame maintains its temperature above the combustion threshold; it vaporizes wax into a continuing supply of fuel; and in usual atmospheric conditions, it induces convection currents, thus pulling in the oxygen it needs and removing the carbon dioxide produced by its own combustion (Bickhard 1998b).

Processes like this tend to maintain themselves; they exhibit *self-maintenance*. That provides another way of expressing this disjunction: *either relatively persistent far-from-equilibrium systems are self-maintenant, or they are not*. The ability to be self-maintaining is an *emergent causal power* of the organization of the candle flame; it cannot be explained simply as the physical resultant of the causal properties of its distinct constituents. Of course, in one sense its persistence is also dependent upon its external conditions: when the candle flame has burnt all its wax, or it is deprived of oxygen, it ceases to be. But so long as the boundary conditions are fulfilled—so long as its external requirements for fuel and oxygen continue to be satisfied—it continues to contribute to its own persistence. It succeeds in maintaining its own process of burning. The ability of a complex system to do this is a holistic property of the system itself. That is one reason why its being self-maintenant cannot be explained in terms of the causal properties of its constituents.

So long as those processes keep the system operational, it will retain its integrity. But once they break down, either because of some fatal disruption from outside or because of aging, the system immediately begins to disintegrate. Some components will decay faster than others—in a dead body, bones last longer than muscles—but there is no sense in which a living body and the corpse left by its death are the *same* body. The former was an integral, self-maintaining system; the latter is *already* in the process of decomposing. That is implicit in the fact that these systems are far-from-equilibrium. Thus, decay proves to be the inescapable dark side of the processes crucial to this ontology.

6 Recursively self-maintenant systems

A further level of complexity is exhibited by systems that can maintain stability not only within certain ranges of conditions, but also within certain ranges of *changes* of conditions. That is, they can switch to deploying *different* processes depending on conditions they detect in the environment. A relatively simple example is a bacterium which has the ability to control its swimming so that it moves towards an attractant chemical. It seems that in *E. coli*, for example, processes along a network of proteins serve to modulate the frequency of its tumbling motion. When moving up an attractant gradient, the bacterium encounters an attractant concentration that increases with time. In response, the frequency of its tumbling decreases and thus it tends to continue moving up the gradient. If it does not encounter increasing concentrations of the attractant, it keeps alternating periods of tumbling and swimming until it does come across an attractant gradient. As a result, it is able to swim towards a source of the attractant chemical (Alon et al. 1999).[11]

These two kinds of activity—swimming and tumbling—are different ways for the bacterium to act appropriately to its environmental conditions. These two ways of acting are ‘appropriate’ in the sense that each contributes to its self-maintenance in the differing circumstances. The bacterium’s ability to detect chemically attractant gradients, and to respond by switching between its two modes of behaving, means that it thereby maintains its own ability to be self-maintenant; it is able to switch between activating one or other of its self-maintenant processes as the environment changes. That is, by means of its internal activity, it exhibits *recursive self-maintenance*.

For a process to be recursively self-maintenant, it must contain within itself some sort of infrastructure which can make the relevant shifts in the system’s own internal processes. A bacterium can switch between swimming and tumbling because, although it is a single cell, it nevertheless contains internal sub-processes that can be activated in response to what yet other internal sub-processes detect in its environment (specifically to differentials in the attractant levels over time). A switching mechanism is the simplest form of infrastructure that can perform this function. More complex organisms contain more elaborate infrastructure which enables continual adjustment to variations detected in their environmental conditions; this is both more complex and more subtle than simple switching.

A relatively stable and cohesive organization of processes that contains within itself sufficient complexity to work in ways that ensure (within limits) its own viability is an *autonomous* system. For the model, that is the significant difference between a candle flame and a bacterium. The complexity that enables the former to be self-maintenant is not internal to the flame itself; those conditions are provided (principally) by the candle and the atmosphere. A candle flame is therefore not autonomous. By contrast, a bacterium is; it strikingly provides for some of its own viability conditions. All biological organisms contain infrastructure of this sort, which enable them to adjust to environmental variation. More than that, the stable structural relationships which cause the components of the system to bind together are not static, as are the bonds which form rocks. Rather, they are constituted by dynamic relationships that continually *re-create* the system itself. Typically, their constituents are replaced many times over during the life of the organism itself. The integrity of such cohesive systems of processes arises from self-generating, self-reinforcing processes.

---

[11] Slightly larger and more complex organisms, such as paramecia, similarly detect whether they are swimming up sugar gradients, using detectors at both ends of their lozenge-shaped bodies. When the one at its ‘front’ detects a higher amount of dissolved sugar than does the one at the ‘rear’, the paramecium keeps swimming; otherwise, it tumbles.

This concept of an autonomous system, which maintains its own integrity though either internal switching or adjustment, yields the next exclusive and exhaustive disjunction through which the ontological model can be elaborated: *some stable far-from equilibrium process-systems are of kinds which are recursively self-maintenant, while others are not*. Those systems are autonomous which satisfy the former disjunct.

These considerations entail an ontology radically different from that standardly offered by physicalism. Biological systems—including human bodies—are *not* to be understood simply as substantial entities ('things' in the strong sense) whose constituents are cells (smaller things), which are in turn (after a few more reductions) constituted out of fundamental particles. Like candle flames, but in ways that are much more complex and sophisticated, biological systems are necessarily open, organized action systems, in *essential* interactions with their environments. Unlike candle flames, through their internal control of such interactions they are able to maintain their own viability conditions and to control their own reproduction. The ontological consequence is that we cannot say what they *are* without taking those interactions into account.

Now, any recursively self-maintenant system is, in at least a minimal sense, goal-directed. Of course, to *describe* it as 'goal-directed' requires an observer. A bacterium does not *know*, in any sense other than a fanciful projected metaphor, that it is seeking some nourishing chemical. Nevertheless, its characteristic way of switching between swimming and tumbling manifests a directedness, a 'towardness', that can reasonably be described as goal-directedness.

At this stage, speaking of goal-directedness carries no implication of consciousness, let alone self-consciousness, on the part of the system itself. Nevertheless, to call such systems 'goal-directed' is neither question-begging, nor anthropomorphic. This concept can be built up from the simpler concepts of 'flow' and the related concepts of 'regulation' and 'control'. In the kind of process-based metaphysics I am sketching here, *everything* is in motion. The ancient intuition of Heraclitus that "everything flows" is confirmed by contemporary physics. That processes 'flow' is the simplest yet most fundamental thing that can be said about them.

How processes flow depends upon the dynamical influences one can have upon another. These can range along a continuum of increasing constraint from 'none' through random perturbations to one regulating the other, to the strongest relationship, when one controls the other. By *regulation* is meant a relationship between two processes such that the two together come to dynamical (or static) equilibrium even if the second, by itself, would not (for example, the motions within the planetary system). By *control* is meant a relationship wherein the first has a reference condition as a goal and dynamically forces the second to match it as closely as possible. Regulation is the wider, but weaker, relationship.[12] That is, the outcome of the first exerts a selection among the possibilities available to the second process. It modifies *how* the second flows, like a stick in a fast-flowing stream modifies the pattern of the water flow (Christensen and Hooker 2000, p. 11). An interactive system shows itself to be a control structure by its manifest ability to test for whether it is in some relevant state,

---

[12] I am indebted to Cliff Hooker for clarifying the difference between 'regulation' and 'control' in this way.

and to adjust its own sub-systems so as to bring itself into, and maintain itself in, that state. Thereby, it keeps directing itself towards that state as the outcome of its own internal processes. This goal-directedness is what licenses, indeed requires, the use of action verbs to describe the behaviour of such a system. We can, and must, speak of what it is *doing*. When the goals selected by the testing, switching, and directing subsystems of an interactive system contribute to the *continued existence* of the system itself, it is a self-maintaining system.

The emergence of recursively self-maintenant systems justifies the introduction of another highly significant concept: action. There is, of course, a loose sense in which any process can be said to be 'doing' something. But the ways in which these systems enable themselves to persist through changing conditions requires that verb to be used in a much stricter and more precise sense. The model I am unfolding here deliberately and explicitly invokes the language of action in a thoroughly serious sense, and makes no sense without it.

I propose three criteria that are each necessary, and jointly sufficient, to warrant what I will call 'minimal' action-descriptions. There is a stronger notion of action which presupposes that, in addition to the three criteria I am about to mention, a piece of behaviour should only be called an 'action' if it is performed by a self-directed agent (a property introduced below). And there is an even stronger sense which requires reflective appraisal of potential consequences; perhaps only humans perform actions in that sense. But these stronger senses of action build upon that of a minimal action, which satisfy these three basic criteria. To articulate these criteria fully would require more space than I have here, so I will only mention them briefly.[13]

Firstly, to count as a minimal action, a piece of behaviour has to be *goal-directed*. Now, all action involves projection into the future; an agent acts *towards* some end. That is indeed the crucial difference between an action and a mere movement—a movement or a series of movements does not necessarily *aim* at anything. Both involve change over time, but actions require more than that. It is characteristic of an action—any action—that it intrinsically involves what the Greeks called a *telos*, an objective towards which it is directed. The structure of action is essentially teleological. Indeed, actions are typically *identified* in terms of their intrinsic ends.

Secondly, to count as an action, a piece of behaviour has to admit the *possibility of error*. A piece of behaviour that is goal-directed can nevertheless miss its mark. That is, it is possible that the organism might discriminate something in its environment that leads it to initiate a procedure that happens *not* to be appropriate in that environment. When it does so, it has manifestly made a mistake. The appropriateness of the behaviour adopted is a *practical* matter, which is why it can be attributed even to something as primitive as a bacterium. Bacteria will also swim up a saccharine gradient. Why that counts as an *error* is that saccharine does not serve the nourishing function for the bacterium that ingesting the right chemicals does. And organisms more sophisticated than bacteria are liable to a wider range of errors in their actions—precisely because they are capable of performing a wider range of actions.

---

[13] I have a more extended discussion of these criteria in my book *Doing the Truth* (in preparation).

The *third* criterion has already been implicit in my exposition of the model thus far. The first two criteria only make sense if the subject to which the action is ascribed is the entire organism; it is *the bacterium* that seeks nourishing chemicals, and can be fooled by a saccharine solution. And it is a *frog* that flicks its tongue and eats flies, and can be tricked into flicking at pebbles. This logical feature is quite general: it is characteristic of action-descriptions that they are attributable to an agent *as a functional whole*.[14]

In a number of interesting respects, the model makes it plain why this should be so. In a process-based metaphysics, entities are *constituted* as cohesive process systems, held together by internal, dynamic bonds. The operation of those bonds is what brings it about that the process behaves in an *integral* way, individuating an entity from its environment. And it is the way an entity's internal process are organized which determines whether it is able to maintain itself in existence, as an integral functional whole, through its modes of activity. That its complex of processes work in *such* a way is the base in reality of our identifying it as an entity. So, an entity's being a singular whole arises from the specific activities and interactions of its constituent processes. Likewise, *recursive* self-maintenance emerges only in certain *whole* organizations of far-from-equilibrium processes. Only they can be can be goal-directed and possibly err in so doing. If any of the sub-processes of such a system be likewise goal-directed and could err they derive those characteristics from their functions within the whole system. This kind of normative behaviour devolves from the whole process to the sub-processes. It does not build up from lower order to higher order, unless the lower orders are already independently normative.[15] Even so relatively simple an organism as a bacterium makes this clear; it makes no sense to ascribe 'swimming' and 'tumbling' to anything short of the organism as a functional whole. Yet those are the activities which are goal-directed, and may go wrong. On the other hand, a chain of causation that passes *through* an organism, but serves no function in its self-maintaining processes, such as a reflex response to an external stimulus, is not an action that *it* performs.

When some piece of behaviour manifest in a system satisfies these three criteria—when that behaviour is directed towards some goal, when it is possible that the behaviour fails to attain that goal and the error is referenced to the self-maintenant condition constituting the actor, and when that behaviour is such that it has to be attributed to the system as a functional whole—it is appropriately described as an action, at least in the minimal sense. Even such relatively simple organisms as bacteria perform actions in this sense.

My identifying the emergence of minimal action at this stage in no way licences a reversion to an individualistic ontology. Actions of any sort, whether minimal or full-blown, can only occur within an interactive context. What legitimates the concept of action is the recursive self-maintenance of cohesive process systems, for which

---

[14] It is necessary to say "as a functional whole", rather than just "as a whole" without any qualification, because a complex organism could lose various bits and still manage to function as a whole. For instance, a person can lose a leg or a hand and still function effectively as a human being, although with more difficulty.

[15] Although they can participate in still higher level normative emergences, even if they are themselves already normative, as do human persons in societies, for example. See below.

continual interaction with their environments is ontologically necessary. Ever since the renunciation of Aristotelian metaphysics in the seventeenth and eighteenth centuries, the concept of action—and its correlate, agency—has been has been deeply problematic, or, at best, regarded as anomalous. Indeed, the denial of agency has been identified as one of three distinguishing features of analytic philosophy in the twentieth century (Capaldi 1991).[16] But we human beings *are* agents, who render ourselves determinate through what we *do*. Hence, identifying the ontological stage at which interactive systems begin to manifest genuine agency is crucial to developing an adequate understanding of who we are.

That interactive systems did emerge, with this distinctive characteristic of being able to perform actions, is a fact that any plausible ontology must accommodate. It follows that goal-seeking, and orientation to the future, is not peculiar to human mentality; it is not a puzzling anomaly only encountered, as one passes up the evolutionary tree, when one finally arrives at human beings. Nor is our recognition of the goal-directedness of biological organisms just the result of an unfortunate tendency of humans to project their own capacities onto regions of the non-human world. The way in which countless far-from-equilibrium processes maintain themselves in existence by manipulating their internal control processes provides the physico-chemical and biological bases in terms of which their actions can be understood. An ontological framework which can accommodate that understanding provides a significant alternative to the prevailing philosophical orthodoxies.

7 Error-detection

On this account, an organism can *be* in error even though it does not have the *concept* of error and does not *know* that it is in error. The next step by which the model is enriched focuses upon those more complex organisms that *can discover for themselves* that some procedure is erroneous, even though they still are not complex enough to have such a concept. This provides the next significant disjunction for the elaboration of the interactive model: *either recursively self-maintenant systems are able to detect that some action they have performed has been in error, or they do not.*

Consider a frog, sitting on its lily-pad, which regularly feeds by flicking its tongue at flies and other bugs in its vicinity. If this frog flicks at a pebble thrown into the air just above its head it will have done something wrong, which can be discovered to be wrong by the frog itself. It will have a surprise—or at any rate, will experience some discomfort—if it succeeds in catching that pebble with its tongue. Even if the frog should swallow the pebble, it will fail to *eat* it. Once more in a minimal sense (for each step taken in building this model should presume no more than necessary), the frog will detect that it is in error. The error it discovers, however, will not be anything about pebbles or bugs; its discovery will be that this was not, after all, a situation offering

---

[16] Lest the above remarks be misunderstood, I hasten to add that reclaiming scientific legitimacy for the concepts of action and agency does not imply any attempt to resurrect Aristotelian metaphysics. For Aristotle and his followers, particular entities—what the medievals called substances—constitute the primary category of being, and that is inconsistent with the model being developed here.

something good to eat. Its tongue flicking and eating action was not appropriate in those circumstances.

Systems as complex as a frog not only have the ability to detect error; they also have evolved an ability to enact one potential interaction when alternative possible interactions are indicated at the same time. Suppose our frog sees a fly and at roughly the same time sees the shadow caused by a hawk flying by. There is no need to ascribe to the frog the ability to compare fly-representations with hawk-representations in order to explain what happens next. Most likely, it will jump into the water rather than flick its tongue at the fly. All that is required to explain its jumping is that these two potential actions be indicated *to* and *for* the frog by its detecting relevant differences in its environment, and that its internal processes enable it to select (in some sense) between alternative kinds of action as a result.

8 Flexible learning

Here again, there is an important distinction. A frog's 'selecting' to jump into the water, because it has detected a hawk hovering overhead, rather than to flick its tongue at a fly it has also detected, need be no more than its having an ability to discriminate between these two affective stimuli and its having an instinctual tendency (developed through evolution) for an aversive reaction to the larger shadow to be the more dominant. After all, these two types of action have very different implications for the frog's ability to keep on being self-maintaining! It is another matter if an organism is capable of *learning*, through conditioning, that some stimulus indicates favourable outcomes, even though a stimulus of that kind would normally be neutral so far as its instinctual responses go. A psychologist's rat, for example, is able to learn which way to turn at the corners of a maze, and to press the blue bar to get food, rather than the red bar (which yields nothing). This rat has learnt which action to favour. It is not too far-fetched to say that it has learnt to *assign value* to what it discriminates and that it has acquired the ability to evaluate the likely outcomes of alternative potential actions.

So, where an organism has the ability to learn which kinds of action yield rewards, and to select amongst potential actions on the basis of that learning, it seems reasonable to say that in a minimal sense it can 'evaluate' these potential interactions. On that basis, we have to say that a frog does not *choose* to jump into its pond when a hawk is hovering nearby. But a rat can anticipate and evaluate the projected outcomes of certain simple potential actions.

The kind of learning involved here is *practical* knowledge; what is learnt is *how* to achieve significant goals. A rat has no theoretical knowledge of why pressing the blue bar should produce food! The maze has been contrived, of course, by a psychologist. But even in the case of much significant human learning, the connections which explain why choosing a certain course of action tends to be a way to produce good outcomes are often not immediately obvious. Even when we have direct informational pathways for evaluating action, we are often ignorant of the underlying processes which serve as effective signals indicating that some potential action would be appropriate at this time.

The need to select amongst alternative goals, such as between a frog's selecting to eat or to save its life by jumping into the water, is not the only kind of selection that complex organisms need to make. As the concept has been introduced so far, goal-directedness has been narrowly conceived as involving a specified routine of determinate tasks, which tend to bring about a specific outcome. But the more highly developed an organism is, the less there is a unique routine of specific actions that the organism must perform if it is to attain a given end. As organisms become more highly developed, learning becomes increasingly self-directed, and behaviour becomes increasingly flexible. In these instances, most of the normative directedness constraining the organism's behaviour are not uniquely associated with a specific set of tasks. There may be numerous *alternative specifications of task-routines* which have the potential to fulfil those same norms. This is so despite the fact that some would perhaps attain that goal more efficiently and more effectively than would others.

In order to deal with these latter cases, the phenomenon of the directedness of action has to be recognized as broader and as encompassing more than simply performing specific task-specifications uniquely determined by a precise goal. Let us reserve the term "goal-directedness" for *specific* goals, associated with specific task-routines. Then the broader notion of generalized goals can be described as involving *generic norms*, which typically do not determine a unique task specification; an indefinite number of outcomes might satisfy them, and there might be indefinitely many alternative task specifications which could yield one or other of those outcomes.[17]

Generic norms in this sense are holistic; they relate to an open-ended range of possible tasks, all of which might satisfy (more or less well) the viability conditions of the organism. These conditions are what the organism, as a functional whole, must satisfy somehow, if it is to persist as an integral system. Consequently, in order to satisfy its generic norms, it has to *select* amongst an open-ended range of potential actions, all of which are to some extent functionally indicated by what it has detected in its environment. A hungry cheetah, for example, has to select which prey to stalk and chase, taking account of the facts that it can be injured by large and dangerous animals, that different potential prey deploy different flight-and-fight strategies, that it has a limited amount of energy to expend, etc., etc. There are no simple and reliable signals that indicate suitable prey, comparable to the role that carbon dioxide plays for mosquitoes.

To satisfy their generic norms, higher level organisms have to be *flexible agents*. They must have the ability to learn from the outcomes of previous actions which of the potential actions available to them in given situations is more likely to succeed. That is, they must have become adept at evaluating the likely outcomes of alternative potential actions, and be able to adjust their behaviour accordingly. Nor do they simply select some specific task routine and switch into it. Rather, any significant action is likely to involve a continual process of appraising and evaluating, of selecting and adjusting—all of which calls upon their previous experience of which actions are most

---

[17] This highlights one of the many deficiencies of the etiological approach to explaining functionality. Etiological theories explain proper functions though evolutionary selection of task specifications, whereas in higher organisms there is not a unique and determinate set of tasks necessary to the attainment of its goals. For a more detailed discussion, see Christensen and Hooker (2001).

appropriate in which kinds of situation.[18] Flexibility and adaptability thus depend upon learning. For a cheetah, this means that it has to learn, through experience, the many interrelated factors involved in successful hunting, including available cover, stalking distance, prey speed and agility, as well as its own capacities for interaction.

This yields the next significant disjunction in the elaboration of the model: *either those recursively self-maintenant systems which can detect that some action they have performed has been in error are of a kind which is able to learn from the outcome of their actions, or they do not.* Higher organisms are recursively self-maintenant systems that can not only detect error, but can also learn from their mistakes and adjust their behaviour through anticipating the likely outcomes of the potential interactions indicated to them by their environmental differentiations. They are 'flexible learners'.

9 Reflective persons

Human beings are much more complicated than bacteria, mosquitoes, frogs, rats, and cheetahs. Nevertheless, many of our abilities have been developed following the same principles. The dynamic model sketched here has unfolded the underlying processes of our self-development. We cannot say what constitutes a human being just by adverting to the cellular components of our bodies. That complex organization of processes that is me extends from my past, and projects into my future, and reaches outside the envelope of my skin. Like any organism, a human is also a system of *necessarily open* processes. Deprive me of interactive exchanges with my environment, and I will soon die. And then the corpse remaining from my former body immediately begins to rot.

However, much more is involved in the emergence of fully human beings than just the biological evolution of our bodies. Traditionally, this 'something more' was described as the acquisition of a 'soul'. And Descartes famously tried to justify this traditional view by arguing that what I essentially am is 'a thinking thing'. Although Descartes is largely out of fashion these days, his way of setting up the issue of what this 'something more' might be continues to dominate the philosophical scene. Today, it is called "the problem of consciousness", and even virulent anti-Cartesians have accepted that they need to address that problem. But there is no such single problem.

To explain that claim, I have to take a short step back. When an organism detects something relevant in its environment, its doing so serves as an indication of a potential action which it might perform. A frog which detects a small dark shape moving in the air nearby is presented with the possibility of flicking out its tongue, which in turn presents the potential action of eating. Here we have an indication of one potential action indicating a further, but connected, potential action. In still more complex organisms, as Mark Bickhard has argued, there may be vast webs of indications of interactive potentialities. Some of these will indicate the potentialities of still others, should those first interactions be engaged in and proceed as anticipated. What we call the representations of objects are constituted as certain forms of invariance

---

[18] For this reason, the tendency amongst psychologists to divide learning into conditioned responses, on the one hand, and cognitive learning, on the other, is too simple-minded. The learning that feeds into this continual selecting and adjusting of motor action is still practical, not theoretical as the cognitivists would have it, but it is too norm-governed to fit into the causal models of stimulus-response conditioning.

within sub-webs of this overall web (Bickhard 1998a,b). These webs are organized in terms of how some interactive possibilities could be reached via various intermediary interactions. To operate efficiently and effectively, the organism has to update and continuously maintain this web. Parts and aspects of it will change with various interactions in which the organism engages, and other changes will occur whether or not the organism engages in specific interactions. There seems every reason to say—and I cannot think of any good reason to deny—that these webs constitute the organism's *knowledge* of its current environment; they constitute awareness, that is, primary *consciousness*.

Clearly, consciousness in this basic sense is not the exclusive preserve of humans. All animals have it. What then of Descartes' inability to doubt that he is thinking? The very grammar suggests the only sensible answer, which nevertheless eluded Descartes himself—and much of the subsequent debate. Thinking that one is thinking, being aware of being aware, has to be a *second-level* operation. Primary consciousness, which we share with the animals, is simply a contentful flow, an experiential flow, but the only way that the *qualities* of that experiencing could themselves be experienced is if there is a second level of the overall system that is interactively, contentfully, experiencing the awareness level of experiential flow. We have to do here with *reflection.* Such a meta-level of experiencing has in fact evolved; it is a characteristic feature of humans. And there is no intrinsic reason why such iterations of experiencing should stop at the number two. We humans can be aware that we are conscious of primary experiencing, and so on.

Much could be said about how this interactive model of experiencing either deals with, or else dissolves, the issues confusedly bundled together under the label of 'the problem of consciousness'. That, however, would be beyond the scope of this paper (For an elaboration of this, see Bickhard 2005). Still, to conclude this outline of our ontological model with the emergence of self-consciousness would be seriously misleading. To do so would encourage the debate to persist within the Cartesian framework. There is yet another level which interacts crucially with that of reflective persons.

10 Social institutions and groups

The step from candle flames to even the simplest biological organisms introduces recursive loops as the activity of the system as a whole has an effect on the internal operations of its constituent processes. Likewise, the development of our distinctly human abilities crucially interacts with the multifarious activities of human sociality. Our individual abilities and traits are significantly affected by the social contexts, institutions, and cultures into which we are born. These social entities are, of course, made up of individual human beings, but human babies develop into mature reflective persons through being nurtured and inducted into a variety of multi-layered social institutions and groups.

So crucial and significant for human development is this nurturing and induction into sociality that the nice symmetry of the ontological model I have been outlining no longer applies. It is not the case that some reflective persons come together to form social groups, while others do not. While a few rare individuals have chosen to live as hermits, even they could not cut themselves off completely from social interaction. Rather, the emergence of reflective persons and the emergence of social institutions and certain social groups are mutually dependent and interactive.

Despite this inter-dependence, there is a strong case for considering the latter as a distinct ontological level, over and above that of individual persons (Bickhard 2004, 2008). This issue has, of course, been much debated, but the basic principle which has driven the identification of ontological levels is clearly evident here also. These institutions and groups manifest properties and powers which are novel and distinctive, and which the humans who constitute them do not individually manifest. As I used to tell my students, it is the university which admitted them as students, and which might eventually confer on them a degree. As an individual, I cannot perform those actions, not because someone else performs them, but because no individual person does. Even when I was the one whose role it was to sign the relevant documents, that was not something I did as an individual person. I was exercising a *role*, a function that only makes sense, and only has validity, because it derives from the structure and dynamics of the institution.

Even looser and more transient associations of people, such as a football crowd, manifest properties and powers over and above those who participate in them. And individuals will perform actions as part of such a crowd that they would never dream of out of that context. Here is yet another example of ‘downward causation’.

The emergence of the properties and powers of these social entities—and how they in turn affect the behaviour of the people who live within them—are probably even more complicated than the emergence of life-forms from chemical systems. Nevertheless, I venture that the appropriate accounts will follow the logical ‘shape’ of the model outlined here.

11 Conclusion

The model thus constructed is schematically presented in Fig. 1, appended. I submit that it offers a radical alternative to the physicalist orthodoxy currently prevailing amongst academic philosophers. Whereas Aristotle and his medieval followers posited *ousiai*—entities, substances—as the primary category of being, Locke, Newton and their modern followers simply substituted corpuscles, basic particulars, in that role. But while that introduced a major upheaval in *physics*, it signalled no *metaphysical* revolution. This model, however, does propose such a revolution, in insisting that entities are themselves emergent, not fundamental. It calls for *processes* to be taken utterly seriously as the basic ontological category, and *interaction* as ontologically necessary.

This model provides a non-reductive schema for understanding ourselves as both embodied and yet emergent, in essential interaction with our environments. In particular, it provides a way of characterising action as ‘metaphysically deep’, not an ontological embarrassment within an otherwise physicalist world. While no doubt much more work needs to be done to refine and flesh out its various ontological levels, I hope that the overall strategy and structure is both clear and powerful enough to be persuasive.

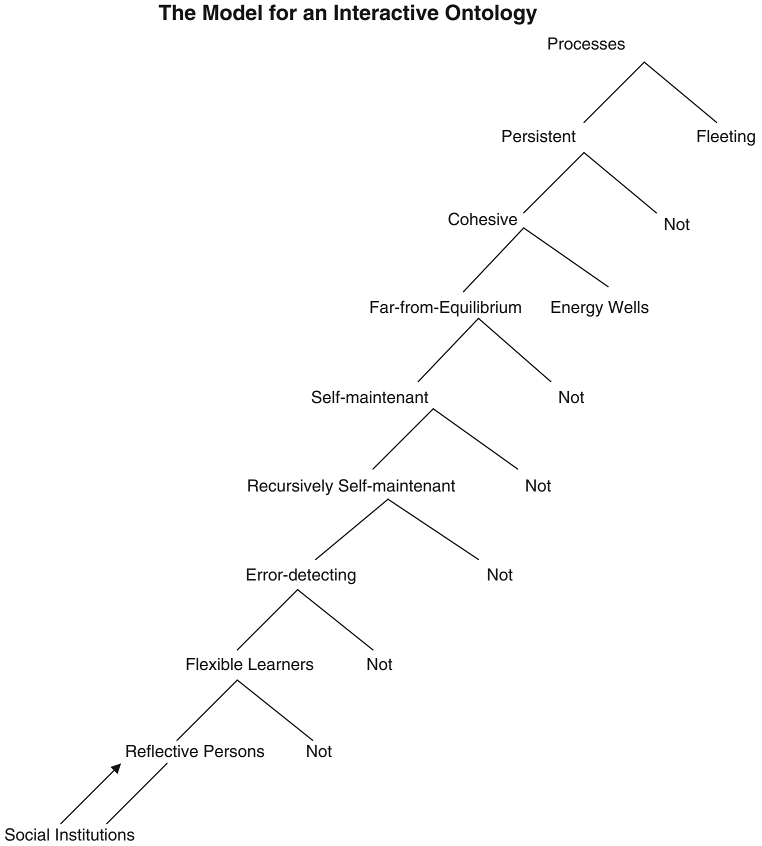

Fig. 1 The model for an interactive ontology

References

Alon, U., Surette, M. G., Barkai, N., & Leibler. S. (1999). Robustness in bacterial chemotaxis. *Nature, 397*(14), 168–171.

Bickhard, M. H. (1998a). Levels of representationality. *Journal of Experimental and Theoretical Artificial Intelligence, 10*, 179–215.

Bickhard, M. H. (1998b). A process model of the emergence of representation. In G. L. Farre & T. Oksala (Eds.), *Emergence, complexity, hierarchy, organization, selected and edited papers from the ECHO III Conference* (pp. 263–270). Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series No. 91, Espoo, Finland, August 3–7.

Bickhard, M. H. (2000). Autonomy, function and representation. *Communication and Cognition, Special issue on Artificial Intelligence, 17*, 111–131.

Bickhard, M. H. (2004). The social ontology of persons. In J. I. M. Carpenter & U. Muller (Eds.), *Social interaction and the development of knowledge* (pp. 111–132). Mahwah, NJ: Erlbaum.

Bickhard, M. H. (2005). Consciousness and reflective consciousness. *Philosophical Psychology, 18*(2), 205–218.

Bickhard, M. H. (2008). *Are you social? The ontological and developmental emergence of the person*. In U. Müller, J. I. M. Carpendale, N. Budwig, & B. Sokol (Eds.), *Social life and social knowledge* (pp. 17–42). New York: Taylor & Francis.

Capaldi, N. (1991). Analytic philosophy and language. In R. Harré & R. Harris (Eds.), *Linguistics and philosophy* (p. 47). Oxford: Pergamon Press.

Christensen, W. D., & Bickhard, M. H. (2002). The process dynamics of normative function. *The Monist, 85*(1), 3–28.

Christensen, W. D., & Hooker, C. A. (2000). An interactivist-constructivist approach to intelligence: Self-directed anticipative learning. *Philosophical Psychology, 13*, 11.

Christensen, W. D., & Hooker, C. A. (2001). Self-directed agenta. In J. McIntosh (Ed.), *Naturalism, evolution and intentionality, Canadian Journal of Philosophy* (Supp. Vol., pp. 19–52).

Collier, J. D. (1998). Supervenience and reduction in biological hierarchies. *Canadian Journal of Philosophy* (Supp. Vol). *Biology and Philosophy*. Calgary: University of Calgary Press.

Davidson, D. (1980). The logical form of action sentences. In *essays on actions and events*. Oxford: Clarendon Press.

Humphreys, P. (1996). Aspects of emergence. *Philosophical Topics, 24*, 53–70.

Jackson, F. (1998). *From metaphysics to ethics* (p. 6, fn. 5). Oxford: Clarendon Press.

Kim, J. (2002). Précis of *Mind in a physical world*. In *Philosophy and phenomenological research* (Vol. LXV, pp. 640–643).

Kim, J. (2005). *Physicalism, or something near enough* (p. 7). Princeton: Princeton University Press.

Lewis, D. (1986a). Humean supervenience. In *Philosophical papers* (Vol. 2). Oxford: Oxford University Press.

Lewis, D. (1986b). *On the plurality of worlds*. Oxford: Blackwell.

Locke, J. (1975). Ed. P. H. Nidditch, *An essay concerning human understanding* (4th ed. of 1700). Oxford: Clarendon Press.

Loewer, B. (2001). From physics to physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and its discontents* (p. 37). Cambridge: Cambridge University Press.

Mourelatos Alexander, P. D. (1978). Events, processes and states. *Linguistics and Philosophy, 2*, 415–434.

Roeper, P. (1987). Principles of abstraction for events and processes. *Journal of Philosophical Logic, 16*, 273–307.

Seibt, J. (1996). The myth of substance and the fallacy of misplaced concreteness. *Acta Analytica, 15*, 119–139.

Seibt, J. (1997). Existence in time: From substance to process. In J. Faye, U. Scheffler, & M. Urs (Eds.), *Perspectives on time, boston studies in philosophy of science* (pp. 143–182). Dordrecht: Kluwer.

Seibt, J. (2001). Formal process ontology. In *FOIS-2001 Proceedings*. ACM Digital Publications.

Seibt, J. (2003). Free process theory: Towards a typology of occurrings. In J. Seibt (Ed.), *Process theories: Crossdisciplinary studies in dynamic categories*. Dordrecht: Kluwer.

Taylor, B. (1997). Tense and continuity. *Linguistics and Philosophy, 1*, 199–220.

Wimsatt, W. C. (1986). Forms of aggregativity. In A. Donagan, A. Perovich, & M. Wedin (Eds.), *Human nature and human knowledge* (pp. 259–291). Dordrecht: Reidel.