A flexible, sloppy blob? Ontology, AI, and the role of metaphysics

Don Ross

University College Cork; University of Cape Town; Georgia State University

don.ross931@gmail.com

(forthcoming in *American Philosophical Quarterly* 2023)

Abstract

Ladyman and Ross (2007) argue that analytic metaphysics is a misguided enterprise that should give way to a naturalized metaphysics that aims to reconcile everyday and special-scientific ontologies with fundamental physics as the authoritative source of knowledge on the general structure of the universe. Le Bihan and Barton (2019) argue, as against this, that analytic metaphysics remains useful as a basis for the body of work in AI known as 'applied ontology'. They stop short of claiming, however, that analytic metaphysics is useful *as metaphysics*. I consider a basis for making the stronger claim: Smith's (1996) project for building what he claims to be metaphysical foundations for applied ontology (and for AI generally). Ultimately, the stronger claim is rejected; but in the course of this dialectic new aspects of the naturalistic metaphysical project come to light, including relationships between it and the traditional metaphysical project of providing foundations for philosophical semantics of truth and reference.

1. Introduction

Humans, along with many other animals, naturally model reality by tracking the movements and transformations of objects. It is consequently no surprise that a main tradition in metaphysics has taken as its core business the elaboration of the alleged general character of objects – what differentiates them, what binds them, what structures them as types and tokens, what transforms, preserves, and destroys them. Extending this perspective, most scientific representation has been, and continues to be, about objects and about two ideas constructed derivatively by reference to objects, properties and relations.

In this context it has been a discomfiting shock to discover that the parts of physics regarded as most general in scope – the parts often called 'fundamental' – ultimately flushed objects out of their basic generalizations (or, as I will say, their laws[1]).

[1] Ladyman and Ross (2007), give reasons for not representing all fundamental physical generalizations as laws. One basis for caution is that laws are often taken to be universal in scope, but it may be that no physical generalizations hold forever (Smolin 1997). Another basis for caution is that perhaps a law necessarily describes processes that evolve in time, as Maudlin (2007, p. 172) argues, but then perhaps (pacé Maudlin) fundamental physical generalizations do not *generally* describe such processes. Neither of these scruples will be relevant in the present essay. It will be

Particles are the most obvious quasi-objects of contemporary physics, but particles, especially in their object-like characteristics, are not the fundamental elements in the ontology. For one thing, in being subject to quantum entanglement particles are shorn of the most basic characteristic of an object according to the Western metaphysical tradition, their individualityacross measurement contexts (French 1989, 1998, 2014; French and Redhead 1988; Ladyman and Ross 2007, pp. 132-145). In addition, particles are *constructed,* as excitations of fields, in contemporary physical theory; they are not primitive kinds. What appear to be the 'brute' elements of physical structure are fields; and fields are radically unlike objects as these have been historically conceived. Whereas individuality in philosophical discourse refers to a technical concept that has a precise analytic meaning (though details of the analysis are continually refined and disputed; see, e.g., Strawson 1959 for a classic discussion), object-hood is a notion in general use in scientific and quotidian contexts, which philosophers should not claim to 'own'. Thus a philosopher who has been convinced that quantum particles are not individuals can always adjust her interpretation of objects in such a way that fields are made to count as objects, and can urge others to endorse this interpretation. But this would amount to simply *forcing* traditional metaphysics into alignment with contemporary physics, and begs all questions about the extent to which physics has or has not *vindicated* traditional metaphysics. An alternative response is to say, with Ladyman and Ross (2007), that traditional metaphysics has been refuted, at least to the extent that its subject matter coincides with that of fundamental physics.[2]

Many philosophers, however, maintain that the extent to which traditional metaphysics and fundamental physics share the same subject matter is limited. According to such philosophers, the implications of Ladyman and Ross's criticism of traditional metaphyscs is at best quite modest: metaphysicians should tread carefully when they venture into provinces abutting philosophy of physics (and physics), and they should add some careful qualifications about the scope of hallowed metaphysical principles some of them defend, such as identity of indiscenibles.

A common platform from which this kind of response has been offered is anti-reductionism. Why should traditional metaphysics be presumed to have the same,

---

taken as important to a generalization's status as fundamental that scientists treat it as having modal force, as Ladyman and Ross (2007) stress. Since this is the feature of lawhood that is essential to any use of the term, I will allow myself to speak here of laws.

[2] Ladyman and Ross (2007) argue at length that if there are no individuals in fundamental physics then there are no individuals at all. This is not crazily presumptuous because individuality, as a technical concept, can be disposed of for technical reasons. Ladyman and Ross do not deny that there are objects (stones, tables, chairs, particle accelerators), and provide an interpretation of them that is intended to reconcile extra-philosophical usage with the absence of objects in the laws of fundamental physics and the denial that there are any individuals.

or a massively overlapping, domain as fundamental physics unless fundamental physics implies an account of everything there is (or would imply such an account to hypothetical beings who could work out the full network of implications of the physical facts)? The thesis that all ontology reduces to the ontology of fundamental physics is embraced by few contemporary philosophers.[3] Among those who reject such reductionism on the basis of extended arguments are Ladyman and Ross themselves (2007, chapters 4 and 5). According to them, the reason that metaphysicians should be particularly constrained to ground their work in fundamental physics neither incorporates nor implies physicalistic reduction. The basis for Ladyman and Ross's 'Primacy of Physics Constraint' (PPC)[4] is instead the fact that fundamental physics is the only part of science - this being close to a definition of what Ladyman and Ross mean by 'fundamental' – that is fully *general*, in the sense of applying with modal force throughout observable reality. (The PPC is not a metaphysical claim, or derived from any metaphysical doctrines such as physicalism or supervenience, but a claim about the institutional and intellectual organization of science.)

This suggests an alternative basis (alternative, that is, to concerns about reductionism) for minimizing the impact of Ladyman and Ross's claims about conflicts between traditional metaphysics and fundamental physics: one might reject their view that it is essential to the mission of metaphysics that it accurately characterize the structure of the world at its most general scale. French and McKenzie (2012, 2015) explore the idea that at least some traditional metaphysics has heuristic value for philosophers of science and scientists, who can "raid" its archive for concepts and identifications of general structural relationships that they can usefully apply in unintended, specific, domains. French and McKenzie recognize that this is a somewhat tepid rescue of traditional metaphysics, and express less than complete conviction that it works. But a less hesitant extension of their kind of response is given by Le Bihan and Barton (2019). They argue that traditional – specifically, analytic[5] – metaphysics has valuable *direct* application to a new branch

[3] A notable exception is Rosenberg (2011).

[4] The PPC states: "Special science hypotheses that conflict with fundamental physics, or such consensus as there is in fundamental physics, should be rejected for that reason alone. Fundamental physical hypotheses are not symmetrically hostage to the conclusions of the special sciences."

[5] Analytic metaphysics is the metaphysical methodology that Ladyman and Ross explicitly reject. They exclude from their criticisms participants in the project associated with Heidegger and other metaphysical phenomenologists (2007, p. 5). Therefore, I will henceforth stop speaking of 'traditional' metaphysics and confine the topic of the discussion to analytic metaphysics, which can be loosely but effectively delineated as philosophy that extends the form of inquiry associated with the work of David Lewis and Saul Kripke. Ladyman and Ross (2007, Chapter 1) provide an extended characterization of the target, but do not attempt to literally *define* it; analytic metaphysics is a historical tradition with vague boundaries, not a doctrine.

of science. The branch in question is the work in information systems and computer science that aims at building *effective ontologies* for deep learning algorithms and for artificial intelligence (AI) representations and interfaces applied to special domains, such as biomedicine. Le Bihan and Barton argue that metaphysics based on contemporary fundamental physics is often (usually) computationally impractical for such applications, and cite specific instances in which information scientists have derived direct inspiration from work in analytic metaphysics that ignores and even in some instances contradicts laws of quantum physics.

This paper critically considers, by an indirect strategy, Le Bihan and Barton's proposal. I call the strategy 'indirect', because instead of engaging closely with their paper I go 'underneath' it by assessing the most explicit effort to find foundations for 'applied metaphysics', the project of Brian Cantwell Smith (1996). The paper is structured as follows. In section 2 I explain why Le Bihan and Barton, by their own admission, stop short of claiming that applied ontology is itself an alternative form of naturalistic metaphysics, even though they express sympathy for that view. In Section 3 I review those aspects of Smith's project that most directly support a metaphysical interpretation of applied ontology. In Section 4 I critically assess this interpretation, partly by comparing it with another project, that of Horgan and Potrč (2008) that reaches similar and complementary conclusions to Smith's, but is grounded in the problems and literature of mainstream analytic metaphysics. Section 5 concludes, arguing that even if one is impressed by the insights of Smith and Horgan and Potrč (as I am), their approaches implicitly *close* the work that remains for metaphysicians to do. Ladyman and Ross, by contrast, leave naturalistic metaphysicians with a busy and open-ended docket of future work. Thus I reject Le Bihan and Barton's suggestion that review of applied ontology shows that analytic metaphysics adds usefully to knowledge. But the more important take-away is that showing *why* this suggestion fails sheds new light on the potential value of Ladyman and Ross's naturalistic program.

2. Applied ontology

From early in the history of AI, it was evident that autonomous computers, as well as cognitively enslaved ones that were nevertheless intended to learn, could not efficiently begin from blank slates with respect to representational categories for objects and processes. A classic instance was Hayes's (1979) Naïve Physics Project, which involved loading a computer with a database of axioms for generating standard condition / action sets to license inference about both manifest relationships amongst mid-scale mechanical types and processes, and consequences of an agent's manipulations of these relationships. This work fostered a wide range of extensions, representatively sampled in Bobrow (1985). Researchers in this line typically described the lore they were representing as 'folk physics'; I can find no published contemporaneous instance in which it is referred to as 'metaphysics', though authors occasionally cite Aristotle's *Physics* and *Metaphysics* as forerunner texts.

Le Bihan and Barton (2019) base their argument for the utility of analytic metaphysics on more recent work in the foundations of computer science, particularly the Basic Formal Ontology (BFO) project of Arp, Smith, and Spear (2015). A principal aim of this specific instance of 'applied ontology' is to facilitate consilience of diagnostic search and inference across different AI systems that all operate on biomedical data. This requires specifications at deep levels of generality, and so, as Le Bihan and Barton illustrate, BFO defines such categories as 'occurrent' (entities that persist over time by maintaining structural relationships among their parts), 'continuant' (entities that are derivative on other entities, such as qualities and dispositions), 'universal', and 'particular'. To support their criticism of Ladyman and Ross, Le Behan and Barton invoke the existence of citations to analytic metaphysicians in the BFO-inspired literature, though they do not provide specific examples.[6]

Let us grant, for purposes of argument, that when designers of semantic architectures for domain-restricted AI need to widen category space to maximize intertranslatability and breadth of inference, they can often find rigorously honed templates in the analytic metaphysics literature. A question that can still be asked is whether what is applied in these cases is metaphysics *as metaphysics*, or whether what is instead going on is that work that was *intended as* metaphysics by its authors is exapted to serve a non-metaphysical purpose. There is arguably an important difference between what Le Bihan and Barton's applied ontologists do with their metaphysical sources and what French and MacKenzie suggest to potential raiders of the metaphysical archive from the philosophy of science. French and McKenzie support the project of naturalized metaphysics, and such metaphysics is viewed by them, as it is by Ladyman and Ross, as extended philosophy of science. Thus French and McKenzie's raiders, when and if they find useful tools to bring home, are still using these tools to do metaphysics, even though the metaphysics in question is not analytic and pays no homage to substantive traditional metaphysical principles, assumptions, or doctrines. By contrast, the applied ontologists aim to develop effective software. It is not obvious that this constitutes doing metaphysics in *any* sense.

The basic reason for doubting that applied ontology is metaphysics is that the underlying aim is not a system of categories and structures that characterizes the *actual* deep structure of the world, but a system that allows machines equipped with

[6] They do cite references to analytic philosophy texts in work around the DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) project of Gangemi et al (2002). This project is focused on reasoning interfaces using natural language, rather than direct engagement between artificial agents and the exogenous environment, so, insofar as analytic metaphysics is regarded as a refinement of folk metaphysics, as Ladyman and Ross (2007) view it, there is whiff of circularity with this example: the metaphysics is useful to the artificial agents simply because their task is communicating with people who structure their thought according to the metaphysics.

it to share conceptual space with human users. Arguably, this enterprise is a species of *cognitive anthropology*. Metaphysics, one might say, aims to accurately describe nature (van Fraassen 2002, pp. 16-17), whereas applied ontology builds extensions of the metaphysical conjectures natural to humans (and then shaped in particular ways by their varying cultures).

Someone might attempt a Kantian move here, and mount a transcendental argument to the effect that it is only possible for people to apprehend the general structure of the world (or, for that matter, any structures) using the deep conceptual system that people in fact operate. Applied ontology, on such an account, might be contended to be an ongoing effectiveness test for such Kantian transcendental analysis. But then, we might be reminded, Kant himself contrasted his transcendental analyses with metaphysics, which he proclaimed as impossible to carry out successfully. If many think that Kant was here quibbling over the meaning of 'metaphysics', and that his achievement is better regarded as showing philosophers how to do metaphysics *properly*, the basis for this is that Kant also thought that his transcendental analysis was uniquely determined by universal principles of reason. There is no hint of any such claim in the work of the applied ontologists, where it would be even more obviously implausible than in Kant's case.[7]

This is not necessarily a problem for Le Bihan and Barton's argument. They do not purport to defend analytic metaphysics *as metaphysics*, and are explicit that their more modest objective is to rebut Ladyman and Ross's claim that analytic metaphysics is epistemically pointless activity. For this objective, it suffices if Le Bihan and Barton can convince readers that work in analytic metaphysics can be exapted to other purposes that are generally agreed to be potentially valuable.[8] But they express sympathy for the bolder hypothesis that applied ontology *is* useful *metaphysical* activity. They defend the plausibility of such a stronger defense of analytic metaphysics (though without mounting it themselves) by arguing that applied ontology is to 'pure' ontological inquiry as applied mathematics is to pure mathematics, and then relying on general agreement that applied mathematics is still mathematics.

---

[7] The moment of reckoning for Kant's interpretation of course came when supposedly uniquely rational geometry turned out not to apply to the real structure of space and time as characterized by general relativity. Quantum physics, the argument goes on, has set fire to most of the rest of Kant's system as a uniquely rational account of fundamental nature.

[8] It should be pointed out that this is a very weak defense of analytic metaphysics, since almost anything that was hard to construct can be used as a source of inspiration. And it implies that all analytic metaphysics should be revisited to render it more useful for AI and related technology, since we no longer think of it as metaphysics. Finally, we should bear in mind that the folk physics project, which was obviously carried out in this spirit, has been largely displaced in the environment of connectionist AI and deep learning.

This starting point for a stronger defense of analytic metaphysics merits close attention,[9] but here I will set it aside in favor of considering a more direct case from the broader AI literature. I thus critically extend the first line of reasoning in Le Bihan and Barton's paper, about applied ontology, instead of following them down their second line.

3. The construction of objects: computer science as metaphysics?

In an idiosyncratic and deeply thought book from 1996, AI researcher Brian Cantwell Smith innovatively explores the following problem space. A system is not autonomous, that is, does not implement a mind that forms, applies and manipulates genuine representations,[10] to the extent that its semantics is rigidly restricted to a fixed domain of objects. Thus the BFO implementation system reviewed in detail by Le Bihan and Barton (2019) is not autonomous because it can operate only within the semantic domain of biomedicine. An autonomous system, Smith contends, must be able to construct and reconstruct its own ontology. But how is the AI designer to set about building such a capacity? Clearly no constructed ontology would be likely to promote cross-domain pragmatic effectiveness if it were not constrained by a semantic link to an external world. But that external world must manifest itself in some format or other. If this starting or basic format is already a domain of objects, then autonomy and genuine representation are, according to Smith's argument, compromised, because, objects are constructed by people in their pragmatic interactions with the world, not simply found.

Dennett (1984; 1987, chapter 8) emphasizes that an individual human does not work up her own ontology from a minimal platform. A person's pre-history of natural selection, and the culture in which she learns, load her with a rich ontology. Furthermore, humans and other mammals probably have innate dispositions to track particular classes of objects through space and time. But these points do not address the motivations behind Smith's project, which at no point appeal to biological verisimilitude in AI. His argument is that the closer to an ontologically minimal platform the AI designer can get, the less likely she is to make what Smith calls 'inscription errors', that is assumptions that because *the designer* can interpret a state of the system as representing a type or token of an object, the system in fact

---

[9] At least two strands of counter-argument warrant careful pursuit. First, one might try to undermine the analogy by arguing that applied mathematics is *not* mathematics, strictly speaking, but physics, or economics, or … . Second, and more fruitfully in my view, one might aim to undermine the analogy by asking whether pure metaphysics enjoys clear principles of sound practice that resemble those that constrain pure mathematicians. But following this thread would lead quickly into deep waters in the philosophy of mathematics, which are not profitably entered for anything but long immersions.

[10] Smith argues for, and does not just assume, that minds require representations. He provides such argument in recognition of the views of cognitive roboticists such as Brooks (1991) who think the embodiment makes representation unnecessary.

uses the representation in question *as* that representation, with all its semantic and effective implications.[11] For example, a face recognition system based on deep learning *might* construct its own concept of face from detecting patterns in transduced color, contrast, and shadow. But most working face recognizers do no such thing; the concept of face is, as it were, forced on them by constraints wired into their programs. We say that such a system 'responds to faces', but the only concept of face in the scenario resides with the user. It is an inscription error to confuse the second sort of case with the first sort. Inscription errors, as Smith says, mistake labeling for empirical discovery. An AI researcher minimizes the extent and likelihood of inscription errors to the extent that her system is built to construct its own objects.

This concern, on Smith's argument, is what connects the foundations of computer science with metaphysics. He argues that the conditions under which it is possible for a mind to construct objects – and on which there is any value in such construction – are stringent and broadly identifiable. To identify them is to identify the general basis of objects. That is surely, Smith contends, to engage in the core metaphysical project. And that is the ambition he takes up in his book.

I will sketch Smith's metaphysical account in broad strokes. My aim here is not to assess its persuasiveness, cogency, or potential for further development. I will instead say just enough about it to provide the reader with a sense of what *kind* of product emerges when metaphysics is seriously conceptualized in this way. The project is self-evidently naturalistic. In Smith's hands, it is extremely interesting, and enlightening in many respects that I will not have space to mention. My aim here is confined to contextualizing Smith's work in such a way as to shed light on the general relationship between naturalistic metaphysics and applied ontology.

Smith, as indicated above, abjures a naïve realist perspective according to which objects exist prior to and independently of representations, and the task of an artificial (or natural) mind is simply to identify them. But neither is his project *radically* constructivist: though objects are constructed by representational systems on his account, they are constructions *for* navigating practically and epistemically in a pre-existing world. If naïve realism were correct, then the obvious place to look for primordial or basic objects would be in physics. Smith argues, however, that although all of physics is ultimately descriptive of the dynamics of particular, material events, physical theory makes no reference to types or tokens of *individuals*, that is, entities distinguished by inherent properties that persist as such through time. This aligns with Ladyman and Ross's (2007) view of the absence of individuals, as well as objects, in physics. However, neither quantum entanglement nor quantum statistics is among the premises of Smith's argument, according to which individuals are already absent from *classical* physical theory. The laws of physics, according to Smith and Ladyman and Ross alike, are verified and tested by application to

[11] Dennett has also insisted on this point, using different language from Smith; see Dennett (1978) chapter 7.

locations that are picked out by concrete material interactions that are specific and pre-theoretical. But whereas Ladyman and Ross are agnostic on the question of how the laws of physics are distinguished from the larger universe of mathematical structures,[12] Smith proposes a semantic demarcation principle: physical generalizations, but not purely mathematical ones, are understood partly by means of deictic application conventions[13] that allow their users to attach them to material instances indexed to idiosyncratic reference frames.

This proposal allows Smith to pursue a subtle strategy for describing the construction of entities that, as a matter of logic, cannot be referred to prior to their construction. First he explores the necessary conditions for object construction that would obtain *if* we could take fundamental physics as already read. Then he characterizes metaphysical construction by analogical extension to contexts in which the laws of physics are *not* relied upon. Thus our knowledge of the laws of physics is used as scaffolding for understanding metaphysics; then the scaffolding is taken down.

Smith makes essential use of two concepts of his own invention: *flex* and *slop*. Before describing these ideas, I note that they appear to be very close to the kind of metric-independent variance of practical application conditions for referential terms, grounded in engineering practice, to which Mark Wilson (2006) appeals in his account of the nature and limits of Duhemian underdetermination of scientific theory. Though Wilson does not cite Smith, their projects and strategies are deeply similar: both seek insight into the preconditions and *value* of abstract, digitalized, theoretical construction out of a pre-theorized, undigitalized plenum engaged with by agents with practical purposes. Wilson characterizes 'facades' as loose summaries of practical engagement histories that become the pre-scientific targets of axiomatic theory construction. Philosophical problems characteristically open in this space because facades are much more adjustable to shifts in practice – have much more flex – than axiomatized theories allow. Thus, according to Wilson, underdetermination is an essential feature of scientific theorizing, not a bug; but it does not have the skeptical consequences for ontology that Duhem (and later Quine) supposed because it is strongly constrained by engineering. Wilson's facades seem to play precisely the role of what Smith calls 'registrations' in the context of his project: arrangements of histories of material contact experiences that are precursors to fully represented objects.

The last point warrants brief elaboration. Smith does not propose or endorse any specific theory of representation. But he assumes, with the overwhelming majority of philosophers, that 'full blown' representation is activity of a *mind*. His focus is on

---

[12] For instance, as Deutsch (2011) argues, the domain of mathematics includes transinfinities, but these have no physical manifestation.

[13] That is, established and generally understood practices for limiting possible perspectives from which the instance can be located. This is also a significant theme in Ladyman and Ross (2007), though they do not use the concept of deixis.

pre-representational processes that must characterize a proto-mind under development by an AI engineer. The finished product, supposing it is an artificial mind, does not *dispense* with these pre-representations, since they remain essential stages of specific representational processing. Similarly, Wilson's facades are used by engineers who have human minds; but before they are artificially stabilized (that is, regimented) in theories, they are not objects of scientific representation in the full technical sense.

Registrations, then, are proto-representational or pre-representational tools for practical engagement with a pre-theorized world. To help us understand the conditions under which they are useful to a system, Smith asks us to imagine two limiting counterfactual kinds of world that are both compatible with the laws of physics.

First, imagine a 'tight' world in which any change of physical state variables at any location propagates instantaneously through all of spacetime. The metaphor here is a set of interlocked rigid gears with no spaces between any adjoining ones. There is in fact an important scientific instance of such a construction: the general equilibrium economy of Arrow and Debreu (1954). In such a deterministic and fully connected economy, no buyer or seller has any decisions to make: all quantities and prices adjust automatically to any event that is exogenous to the system. That is to say, in Smith's terms, that the system has no *flex*. Actual, empirical, markets, by contrast, do exhibit flex, because shifts propagate with diminishing extent as we move outward in economic space (which often, though not always, coincides with physical and political geography): if the pizzeria on my street in Cork, Ireland, raises its prices in response to an exogenous local factor, this will influence the supply and demand schedules of fast food throughout my neighbourhood, prices of clothes in Dublin only by indetectably small margins, and prices in Cape Town not at all. Thus investors in actual markets, unlike hypothetical investors in a market in perfect general equilibrium, need to distinguish pizza from clothes, Cork from Dublin, and Ireland from South Africa. Smith argues that there is no use for registrations or (therefore) represented objects in a tight world with no flex. The field equations governing the whole system tell us everything it is useful to know.

Next, consider the opposite limiting case, in which the world has no metaphorical gears at all: all physical magnitudes covary (according to the laws of physics) smoothly, continuously, and uniformly, with the laws applicable purely locally at every scale. Here, flex is limitless: there is no distinction between causes and effects, so no potential value in distinguishing any registration from any other, no need for anything but laws and locations, so no potential value in constructing and representing objects.

Slop refers to the tendency in the actual world for flex to vary inconsistently with the distributions of the state variables specified by the laws. To return to the analogy from economics, the price of pizza in Cork doesn't influence the price of pizza (or anything) in Cape Town, but the price of very expensive seaside houses in Cork *does* influence the prices of similar homes in Cape Town; the segmentation of

markets features cross-cutting dimensions. Smith is not explicit about whether slop is *another* necessary condition for registration and object-construction over and above the need for flex; he simply argues that mid-range levels (i.e., levels away from the limits) of 'flex and slop' are necessary.

Agents that register local variations allowed by flex and slop can gain forecasting leverage, and therefore a measure of effective control, if they can track unevenly distributed covariations at higher orders of abstraction. For this they need to partly *registrationally disconnect* from the scale of immediate presence at which they are materially engaged with complete realizations of concrete events. This has another important payoff: they can model, by inference, events with which they become *materially* disconnected: an object maintaining its trajectory while it passes behind a barrier, the career of Napoleon Bonaparte two centuries ago, or the prices of houses on which one isn't currently in the act of bidding. This is the basis of representation.

As perspectives become more abstract, increasing details of information about concrete particulars are discarded, or at least not attended to.[14] In the scaffolded argument space considered so far, the ultimate scale of abstraction is fundamental physics, where, according to Smith, the laws can be interpreted as applying to any material particulars at all, even though, to be understood, they must be associated with at least hypothetical particular material states.[15] If physicalistic reductionism held, then this ultimate abstraction would 'contain' the details of all less abstract representations. If 'contained' means 'imply by an effective computation' then this would indicate physical determinism at all scales; if 'contains' means only that all neglected details are recoverable as the scale of abstraction is reduced, then we would have (global or local) supervenience on physics. However, because, according to Smith, physical theories apply to material interactions without ever *fully* specifying them, and because flex and slop occur non-uniformly across dimensions that physical theory doesn't recognize, reductionism is false.

Philosophers are apt to be unhappy with this dismissal of reductionism by resort to metaphors. But Smith makes no pretense of providing any original argument for anti-reductionism; he takes it as having been established by philosophers of science. What is important for present purposes is that it furnishes Smith's license for regarding metaphysics as a domain of abstraction that is analogous to physics in its relationship to material particulars and locations, but distinct from physics and not implied by it. Smith's account of metaphysical abstraction is metaphorical, unlike his

---

[14] This seems to have roughly the status of a tautology in Smith's analysis.

[15] Thus, as Smith emphasizes, going to fundamental physics does not involve *complete* disconnection from material experience. On Smith's account, the idea of completely disconnected representation is in fact incoherent. He does not indicate what he thinks, in this respect, about mathematical objects that are physically impossible. He might take them to be meta-representations, and thereby necessarily connected to the material world through signs.

account of physical abstraction, because we do not have agreed-upon 'laws of metaphysics'. This is why the bulk of his book is occupied with the scaffolded construction of objects under the constraints of physical laws, even though a system's object construction is held to be ultimately regulated by metaphysical principles that do not derive from physical theory. One such principle identified by Smith is that an object, to be worth constructing, must detectibly persist through some interval of time. (Thus Smith can agree with Ladyman and Ross that some elements of fundamental physics, or perhaps even all elements of fundamental physics, might not be objects.)

What emerges from all this, as the reader will have gathered, is not something that looks like a traditional metaphysical theory. But that is clearly not Smith's objective. What is of interest here is the extended effort to demonstrate what metaphysics is *for* in the context of building a mind that can function effectively in an open-ended world. *Because* the system must cope with flex and slop, according to Smith, it must have an architecture that can consolidate its own history of registrations across multiple dimensions of abstraction, including dimensions that cross-cut the dimensions featuring in physical laws. It will develop an implied logic of objects that can be inferred from its behavior and from the representations that will be characterized in a computer scientist's theory of it. Doing AI research does not directly teach us the true metaphysics of the world, on Smith's account, but it shows us what metaphysical construction is, and is for. This, I contend, is a much more prima facie persuasive candidate for useful application of metaphysics in applied ontology than any of the examples considered by Le Bihan and Barton (2019).

4. Naturalistic uses of metaphysics

The review of Smith's work above gives us grounds for thinking that applied ontology might be genuinely relevant to metaphysics, and not just to cognitive anthropology. But now we must step back and assess the extent to which Smith's project is aligned with Ladyman and Ross's conception of naturalistic metaphysics. Should Smith's work motivate Ladyman and Ross to relax their skepticism about the value of analytic metaphysics in the direction urged by Le Bihan and Barton?

According to Ladyman and Ross (2007), the purpose of metaphysics is to identify senses in which science is unified – that is, the senses in which the scientific disciplines as a body provide a single, coherent model of empirical reality – even given radical failure of reductionism.[16] They argue that given this purpose, every metaphysical argument must attend to details of fundamental physics, because that is the only domain of scientific knowledge that, like metaphysics itself, is taken to constrain all measurements of all dimensions of reality at all scales.

[16] By 'radical failure' I allude to skepticism about even the weakest form of physicalism that endorses global supervenience of the non-physical on the physical. See Ladyman and Ross (2007, pp. 38-57), and Ross (2016).

Must Ladyman and Ross, in light of this, deny the feasibility or coherence of Smith's removal, near the end of his book, of the scaffolding of physical law on which the bulk of his account of the construction of objects depends? The answer is negative: an AI system could be designed in such a way that, as a consequence of its practical engagements with the world, it constructs objects according to constraints that are not derived from, or are even inconsistent with, the laws of physics. This might even be the most effective path for the system to take in its object construction; arguably it is what actual people do insofar as they *construct* objects rather than genetically or culturally inherit them, particularly since what most people take to be laws of physics are in fact just the classical heuristic generalizations that are good enough for mechanical engineering near the surface of the earth. But this brings us directly back to the impasse left by Le Bihan and Barton: even if people draw on traditional metaphysical theory when they construct some of their objects, are they not then just using tools developed by metaphysicians for non-metaphysical purposes? If the answer to that question is affirmative, why should we not say the same thing about what applied ontologists allegedly do when they borrow resources from the archives of *analytic* metaphysics? (It is worth noting in this context that although Smith does cite work by analytic metaphysicians, his actual metaphysical account relies on core concepts, flex and slop, that he forges himself from raw materials in computation theory.)

Perhaps the apparent impasse can be escaped by considering purposes that other philosophers who are at least not hostile to naturalism – even if Ladyman and Ross might, in their stringency, regard their naturalism as compromised – have aimed to serve through metaphysical inquiry. The obvious big enterprise here is explicating the core semantic concepts: truth (by correspondence) and reference. If a philosopher believes that the basic equipment of analytic philosophy, logic and set theory, are reliable machinery for building rigorous accounts of truth and reference, and if they think that this depends on identifying the right kind of correspondence relationship between, on the one side, referential terms and predicates and, on the other side, objects and properties, then they are motivated to seek plausible metaphysical structures that are, as it were, appropriately structured locks for their keys. Ladyman and Ross resist this metaphysical methodology because they interpret the recent history of physics as undermining the capacity of logic and set theory to get an accurate grip on the most general structures identified by science. They look instead to dynamic information theory, and more fundamentally to group theory and to mathematical novelties still being refined, such as homotopy types (Univalent Foundations Program 2013).

To avoid begging questions against Smith, let us for now set aside Ladyman and Ross's grounds for radical revisionism based in their arguments from quantum theory. The closest approximation to Smith's metaphysics in the substance of its conclusions that is still anchored to the project of building foundations for philosophical semantics is arguably the project of Horgan and Potrč (2008).

Like Smith, and Ladyman and Ross, Horgan and Potrč conclude that neither the objects of everyday folk discourse nor the objects referred to in the theories of the special sciences are discoverable in the independent fundamental ontology. However, the premises that drive this conclusion for them are mainly grounded in philosophical semantics: most objects **ϖ** as modeled in everyday discourse and special sciences are vague with respect to their boundaries; but it is a category mistake to suggest that actual, mind-independent objects **ϑ** could exist vaguely. Therefore, the **ϖ** must be constructed. But Horgan and Potrč reject error theory about reference to them; statements such as "There are 8 chairs around the dining room table", "The human brain is larger than the human kidney", and "The wombat is native to Australia" are asserted as true and governed by truth-functional, standard semantics. The mechanism developed by Horgan and Potrč to resolve this dilemma, which they incorporate into a theory they call 'austere realism', is contextual semantics. In contexts where people use 'positing expressions' ("there are … **ϖ**'s"; "∃(x)(x=**ϖ**)"; "the effects would be explained if caused by **ϖ**'s" etc.) to make ontological claims, the governing semantics is *direct correspondence* (DC): each **ϖ** must truly correspond to exactly one **ϑ**. But DC is the governing semantics, in general, only when people are engaged in metaphysics. Otherwise, the governing semantics is *indirect correspondence* (IC): an act of sincere positing commits a speaker only to the claim that the world is structured such that, in the prevailing context of usage, there is (local) warrant for an assertion for which the Tarskian T-sentence includes "**ϖ**" on the left-hand side and mentions **ϑ** on the right-hand side. The model of the activity by which communicating people identify the context-setting parameters is Lewis's (1979) "scorekeeping in a language game". In IC contexts, positing is said to involve *ontic* but not *ontological* commitment.

Horgan and Potrč argue that various metaphysical theories are compatible with austere realism. The one that they defend as most efficient is 'Blobjectivism'. This is the thesis that the set of **ϑ**'s has only one member. As this object has no proper parts (since those would be additional **ϑ**'s), it is a blob. It has enough structure to support the laws of physics, and all the dynamics that give rise to the special sciences and the variety of life, mind, and culture that allow for endless true utterances in contexts governed by IC. But this structure is irreducibly dynamic and complex and doesn't resolve into ontologically stable objects (that is, into objects that can be picked out non-vaguely). One might wheel in Smith's (1996) story here and suggest that the Blobject has just the goldilocks levels of flex and slop to make it useful for cognizers that flicker through it to collectively construct a zoofull of **ϖ**'s.

Both austere realism and Blobjectivism are elegant constructions, and their harmonious amalgamation is still more so.[17] The similarities between the resulting metaphysic and Smith's construction are striking, given the radically different motivations for the projects. (It is important to note that whereas Smith sees the project as providing foundations for a computational model of mind and

[17] In the structural realism literature, French (2014, p. 174) expresses queasiness about relativistic implications for truth in austere realism.

representation, Horgan and Potrč argue that austere realism is incompatible with such a model. The model with which it is argued to be complementary, i.e. standing in a relation of mutual support, is a dynamical connectionist one, in which computable multi-dimensional physical state transitions support cognitive state transitions that are *not* generally computable (Horgan and Tienson 1996). The absence of fully specifiable discrete cognitive states is part of the explanation for the inability of dynamical cognitive systems to refer non-vaguely to any ɵ's except the Blobject.)

But now it is time to ask: what *work* is metaphysics doing in these projects? In Smith's case metaphysics reminds us not to precommit an AI system to any particular objects lest we fall into inscription errors. In the case of Horgan and Potrč, metaphysics relieves us of anxiety arising from the fact that standard semantic analysis and realism about the external world seem to be in *prima facie* conflict. We avoid this by seeing that we can ontically commit ourselves to tables, brains, kidneys, wombats, and Australia while ontologically committing ourselves only to the Blobject. In both cases the work done by metaphysics is minimal – Wittgensteinian philosophical therapy rather than ongoing knowledge production – and it has basically been done and completed for us by the respective authors.

One might aim for a more generous interpretation as follows: Smith's metaphysics shows us how a computational theory of mind and representation can be constrained by physical law without reducing to physics; and austere realism shows us, partly by reference to some serious complex-systems and connectionist cognitive science (without which the account would be quite unsatisfying), how special sciences can have distinctive ontologies from one another while still all emanating from one common blob of a world. We arrive at these generous interpretations by understanding the purpose of metaphysics in the way defended by Ladyman and Ross: explaining how the sciences collectively deliver a coherent general worldview despite the failure of reductionism and perhaps even of global supervenience on the physical.[18]

A crucial difference we can now identify between Ladyman and Ross's project and the others is that the former indicates an *ongoing agenda* of further metaphysical work. If metaphysics is derived directly from fundamental physics instead of being separated from it (Smith) or transcending it (Horgan and Potrč), then any and all new details uncovered by fundamental physicists should be expected to inspire new metaphysical details. Ladyman and Ross's metaphysical program generates a range of separate sub-programs: reconcile details of fundamental physics with details of geology, and then also with details of economics, and then with details of population

[18] Horgan and Potrč implicitly sign up for Ladyman and Ross's naturalistic program, in a way that Smith clearly does not, when they argue that it is a virtue of Blobjectivism that the Blobject resembles a quantum field-theoretic interpretation of the universe in supporting complex dynamics while not having proper parts or containing objects.

ecology, and so on. By contrast, Smith and Horgan and Potrč, while taking metaphysics fully seriously, seem to aim to close the books on it by giving us respective last words.

If Smith, in particular, does metaphysics in order to more securely isolate the practical problem for applied ontologists, then Le Bihan and Barton's suggestion that applied ontologists should look for inspiration in analytic metaphysics seems much less persuasive than it might appear to be at first glance. It isn't really metaphysics that the applied ontologist needs on Smith's computational approach to AI, but semantics. Smith supplies no details about the semantics in question, but Horgan and Potrč offer plenty. Then, however, with our metaphysical *worries* resolved by confidence in the serene existence of the Blobject, AI research can rely on *linguistic* semantics or perhaps more liberally on semiotics (in light of the irreducible dynamics and all that) while setting traditionally *philosophical* semantics aside.

Applied ontology, then, doesn't appear to be applied *metaphysics*. Is there, for that matter, any distinction to be drawn between 'pure' and 'applied' metaphysics? My answer is: no. Unifying science is the *activity* in which a defensibly naturalistic metaphysics consists.

Thus I reject Le Bihan and Barton's suggestion that work in applied ontology should motivate Ladyman and Ross to temper their skepticism about the value of analytic metaphysics to the edifice of human knowledge. If even Smith's self-proclaimed metaphysics for the *foundations* of AI don't really make significant use of analytic metaphysics, or imply that researchers extending his project should read the analytic metaphysics literature for guidance, then its value seems unlikely to be rescued by the more conservative kind of work that Le Bihan and Barton consider. But the discussion has not yielded only a negative conclusion. The cost of trying to make metaphysics 'safe from' the amazing ontological implications of fundamental physics is to make metaphysics *dead*. By contrast, a metaphysics *driven by* fundamental physics, such as Ladyman and Ross urge, has a work program until physics is finished, which likely means, until the whole human intellectual enterprise ends.

References

Arp, R., B. Smith, and A. Spear (2015). *Building Ontologies With Basic Formal Ontology.* MIT Press.

Arrow, K., and G. Debreu (1954). Existence of an equilibrium for a competitive economy. *Econometrica* 22: 265-290.

Bobrow, D., ed. (1985). *Qualitative Reasoning About Physical Systems.* MIT Press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence* 47: 139-159.

Dennett, D. (1978). *Brainstorms*. Bradford.

Dennett, D. (1984). Cognitive wheels: The frame problem of AI. In C. Hookway, ed., *Minds, Machines and Evolution*, pp. 129-151. Cambridge University Press.

Dennett, D. (1987). *The Intentional Stance.* MIT Press.

Deutsch, D. (2011). *The Beginning of Infinity*. Allen Lane.

French, S. (1989). Identity and individuality in classical and quantum physics. *Australasian Journal of Philosophy* 67: 432-446.

French, S. (1998). On the withering away of physical objects. In E. Castellani, ed., *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*, pp. 93-113. Princeton University Press.

French, S. (2014). *The Structure of the World*. Oxford University Press.

French, S., and K. McKenzie (2012). Thinking outside the toolbox: Towards a more productive engagement between metaphysics and philosophy of physics. *European Journal of Analytic Philosophy* 8: 42-59.

French, S., and K. McKenzie (2015). Rethinking outside the toolbox: Reflecting again on the relationship between philosophy of science and metaphysics. *Poznan Studies in the Philosophy of Science and the Humanities* 104.

French, S., and M. Redhead (1988). Quantum physics and the identity of indiscernibles. *British Journal for the Philosophy of Science* 39:233-246.

Gangemi, A., N. Guarino, C. Masolo, A. Oltramari, and L. Schneider (2002). Sweetening ontologies with DOLCE. In *Proceedings of the International Conference on Knowledge Engineering and Knowledge Management: Knowledge and the Semantic Web*, pp. 166-181. Springer.

Hayes, P. (1979). The naïve physics manifesto. In D. Michie, ed., *Expert Systems in the Microelectronic Age*, pp. 242-270. Edinburgh University Press.

Horgan, T., and M. Potrč (2008). *Austere Realism.* MIT Press.

Horgan, T., and J. Tienson (1996). *Connectionism and the Philosophy of Psychology.* MIT Press.

Ladyman, J., and D. Ross (2007). *Every Thing Must Go: Metaphysics Naturalised.* Oxford University Press.

Le Bihan, B., and A. Barton (2019). Analytic metaphysics versus naturalized metaphysics: The relevance of applied ontology. *Erkenntnis*, forthcoming.

Lewis, D. (1979). Score-keeping in a language game. *Journal of Philosophical Logic* 8: 339-359.

Maudlin, T. (2007). *The Metaphysics Within Physics.* Oxford University Press.

Rosenberg, A. (2011). *The Atheist's Guide to Reality.* Norton.

Ross, D. (2016). Vikings or Normans? The radicalism of naturalized metaphysics. *Metaphysica* 17: 213-227.

Smith, B.C. (1996). *On the Origin of Objects.* MIT Press.

Smolin, L. (1997). *The Life of the Cosmos.* Oxford University Press.

Strawson, P. (1959). *Individuals*. Methuen.

Univalent Foundations Program (2013). *Homotopy Type Theory.* https://homotopytypetheory.org/book

van Fraassen, B. (2002). *The Empirical Stance*. Yale University Press.

Wilson, M. (2006). *Wandering Significance.* Oxford University Press.