# Building Up: foundations and material for definitional ontology construction[*]

Brandon Bennett[1]

[1]School of Computer Science, University of Leeds, LS2 9JT, UK

**Abstract**

This paper investigates the role of foundational theories in the construction of formal ontologies. Whereas current practical ontologies employ foundational theories primarily to organise the upper-level categories of their class hierarchy, this paper explores an approach that builds upwards from fundamental concepts of space, time, and matter. The paper explores the foundations of a definitional methodology in which all concepts are firmly grounded upon basic primitives and the vocabulary is extended incrementally towards more human-oriented, mid-level vocabularies. A significant advantage of this approach is that when new vocabulary is introduced, no additional axioms are required beyond the definitions themselves: semantic constraints holding between defined concepts are automatically enforced by the axioms governing the primitive terms. The paper highlights some meta-logical concepts that can help identify definability relationships among terms within an ontology. The gap between foundational theories and more general conceptual terms is acknowledged and ways to bridge this are explored. In particular, the paper considers the nature of *matter* and how this can be introduced into a foundational theory. In addressing this, the paper examines a puzzle, first identified by Tarski, regarding the axiomatisation of a theory of matter and mechanics within any framework that incorporates a theory of geometry. A solution to the puzzle is proposed in which a primitive concept of 'matter' acts as both a symmetry breaking reference frame that conceptually enriches a purely spatial ontology, and also as a representation of the actual material substance of the world.

**Keywords**

Foundational Ontology, Ontology Construction Methodology, Definitions, Ontology of Matter

## 1. Introduction

Since ancient times, a central goal of those investigating the nature of reality has been to determine the fundamental elements of which it is composed, and how the rich and varied structure of the world we observe is constructed from those elements. This goal has been pursued both in relation to the material constituents of the world (culminating in atomic and sub-atomic physics) and also its conceptual elements, the basic ontological categories that underlie our apprehension of the world in thought and language.

Scientific theories have been developed at several different levels of scale, and in relation to different types of element and different modes of behaviour. Hence, we have atomic and sub-atomic physics, chemistry, mechanics, cosmology, biology, psychology etc. But, notwithstanding this division into sub-disciplines, our modern conception of science is that it must regard all aspects of the world as operating according to a single set of fundamental principles, and moreover that structure and behaviour observable at large scales is completely determined by the behaviour of smaller scale elements of reality. Indeed, the interactions between different fields of science (such as between microscopic and macroscopic physics, chemistry and biology, biology and psychology etc.) are areas of investigation that have resulted in theories providing bridges between them (e.g. we have well-developed accounts of

---

*Proceedings of the Joint Ontology Workshops (JOWO) - Episode XI: The Sicilian Summer under the Etna, co-located with the 15th International Conference on Formal Ontology in Information Systems (FOIS 2025), September 8–9, 2025, Catania, Italy*

[*] Thanks to the anonymous FOUST and FOIS-2025 reviewers for valuable comments on earlier versions of this paper.

B.Bennett@leeds.ac.uk (B. Bennett)
https://bb-ai.net (B. Bennett)
0000-0001-5020-6478 (B. Bennett)

how atomic physics explains chemical interactions).

## 1.1. Aristotelian Ontologies

The person most responsible for the conception of *ontology* and who has had a strong and lasting influence on its development is Aristotle.[1] His method of analysis particularly emphasised the *species*, in other words *natural kinds* of things, as being the level of categorisation that most clearly reveals the essential characteristics of things (i.e. the properties that are intrinsic to their nature rather than just a result of particular circumstances).[2] Starting with a loose idea of similar things forming a species, one can define species more precisely by: first, considering broader categories, *genera*, that incorporate several species; and then, identifying *differentia*, which are the salient properties that divide a genus up into its separate species.

The Aristotelian approach investigates and reveals a hierarchical structure, or *taxonomy*, that classifies elements of reality into groups and describes the logical relationships between these groupings and the properties that give rise to these relationships. Certainly, this kind of structure and the logic that it captures are useful for many tasks involving manipulation of information expressed by means of conceptualisations based on those used in human thought and language. Because of this, it plays a core role in the design of modern ontologies that are used for computational processing of information to support human activities. The extent to which modern computation-oriented ontologies are developed according to Aristotelian principles varies. The stated methodology of BFO explicitly acknowledges its Aristotelian influence [2], whereas the designers of DOLCE took a more cognitive perspective that is further from Aristotle's approach [3]. Other foundational ontologies such as GFO and UFO adopt more mixed strategies: GFO combines realist and constructivist elements in a formal framework that partially reflects Aristotelian structures [4], while UFO incorporates both ontological realism and conceptual modelling concerns influenced by cognitive and linguistic considerations [5].

## 1.2. Reductivist Approaches to Ontology

Aristotle's approach to describing reality can be contrasted with that of Democritus (and subsequently Epicurus, who held a similar view). Democritus was more interested in the physical and material properties of the world. He proposed the existence of atomic particles that were the ultimate constituents of reality and considered large-scale physical objects and their behaviour to be completely determined by the properties of their microscopic atomic constituents. It is clear that the Democritean conception of physical reality is closely aligned to, indeed prophetic of, the theories of modern physics. Nevertheless, with regard to conceptual analysis of reality by philosophers through most of history, and right up to present day ontology development, the species-based approach of Aristotle has eclipsed the atom-based approach of Democritus.

But there are exceptions to this. Perhaps most prominent is the methodology of Carnap, as propounded in his influential book *Der logische Aufbau der Welt* [6]. The approach can be described as *definitional* and is based on the idea that the meanings of all terms one may wish to deal with can be fixed in relation to a small number of fundamental concepts and relations. This is achieved by building up layers of definitions that ground the meaning of each concept, either directly upon primitive concepts and entities, or in terms of other concepts that have already been so grounded. This approach was influenced by Russell, who had not only attempted to found mathematics upon fundamental principles but also suggested that our entire conceptualisation of the world can be defined in terms of primitive sense data [7]. More recently, in his paper *The study of ontology* [8], Fine gave a general analysis of the structure of ontologies in terms of complex entities being constructed from simple 'given' elements, and similarly general approaches have been pursued in [9] and [10]. Whereas the Aristotelian approach might be called 'weakly definitional' in that it defines species in terms of their characteristic differentia,

---

[1] The discipline was not explicitly demarcated and named until centuries later by Rodolphus Goclenius [1].

[2] "*Essence will belong to nothing except species of a genus, but to these only; for in these the predicate is not considered to be related to the subject by participation or affection, nor as an accident.*" (Aristotle, Metaphysics, Book VII)

the Russell-Carnap approach is much more strongly definitional in that it seeks to define all concepts in terms of a small number of primitives.

The Carnapian project was extremely influential and was the major force behind the *Logical Positivist* movement in philosophy during the first half of the 20th century. However, by the latter half of that century, its influence had declined due to technical difficulties, lack of obvious applications and the move towards a view of language as being too flexible and pragmatic to be amenable to description in formal logic. Nevertheless, now that the value of ontologies in the fields of computation and information science is widely recognised, and the challenges of constructing such ontologies are accepted as being very large but worth tackling, the possibility of constructing strongly definitional ontologies may be more plausible and attractive.

### 1.3. Reductivist Ontology in Relation to Scientific Reductionism

The reductivist approach to ontology that I am considering is related to but distinct from the notion of reductionism that is often debated by philosophers of science (e.g. [11, 12, 13]). In that arena, the focus is mainly upon the question of whether theories of the higher level sciences (e.g. psychology) are fully explainable in terms of the terminology and laws of lower level sciences (such as chemistry, and ultimately physics). If this kind of reduction is possible, one would expect that: definitional reduction of terms from higher level theories to those of low-level theories is also possible; and that, causal laws need only be formulated in terms of the conceptual vocabulary of the lowest level theory (which would then entail the laws of the higher level theory). But even if such terminological reduction is possible it is unclear and highly contested whether that entails that laws formulated in terms of high-level concepts can always be adequately articulated in terms of lower level concepts (various counter-arguments have been put forward, e.g. in [14, 15, 16, 17]). The current paper does not attempt to enter this debate but my presumption is that although, definitional elimination of high-level terminology would in principle enable high-level laws to be reduced to those in a lower level theory, but with the following caveats: (1) in most cases the resulting expanded versions would be highly intricate, uninformative and intractable; (2) the behaviour of complex systems cannot be neatly factored into a product of behaviours of their parts (because of $n$-body problem and quantum entanglement/coherence); and (3) some primitives may be required that go beyond those apparent in standard formulations of physics (e.g. primitives explicitly referring to some aspect(s) of possibility or observation).

### 1.4. The Role of Foundational Theories

What is the role of foundational theories in the construction of ontologies? The obvious answer would be that they provide a foundation upon which the rest of the ontology is built. Yet it seems that typically this is not the case. In the development and use of practical ontologies that are actually employed within information systems, foundational theories (such as theories of space and time) mostly play a more peripheral, albeit important role. They typically provide top-level categories used to classify entities into distinct ontological kinds and correspondingly distinguish lower-level categories according to the ontological type of their instances. For example, the Cell Ontology and Gene Ontology use BFO [2] in this way. Ontoclean [18] has been widely used as a tool for checking the internal coherence of taxonomic structures. It provides useful integrity checking by applying formal meta-properties (such as rigidity, identity and dependence) to the upper-level categories in an ontology (which will be inherited by their sub-categories), helping to detect and correct classification errors and inconsistencies. Foundational theories, in the form of upper-level ontologies, are also beneficial for ontology integration and alignment. One demonstration of this is the large number of ontologies in the OBO Foundry [19] that are integrated within the BFO framework. The Suggested Upper Merged Ontology (SUMO) [20], although not widely adopted as a standard, was specifically developed to provide a general-purpose framework for integrating domain ontologies through a set of high-level categories and logical relations. Unfortunately, different upper level ontologies (e.g. those of BFO, DOLCE, GFO and UFO [2, 3, 4, 5]) tend to differ in ways that are hard to resolve. The TUpper [21] framework provides a modular collection of upper-level ontologies that can be linked to domain ontologies. Its methodology is orthogonal and potentially complementary to the definitional strategy pursued here.

## 1.5. Do ontologies need to be grounded?

One reason why a grounded approach has not been adopted is that it is not necessary for many applications that are limited to specific domains. Purely taxonomic reasoning is very useful and can be captured by general relations between species without one needing to ground the meanings of species concepts. And when the differentia underpinning the taxonomy are also represented as explicit combinations of properties and relationships, the system becomes more powerful still.

Yet, the logic of taxonomies is only one aspect of human reasoning. To understand a sentence such as: *"Sue often played with her friend Tom, but Tom's father owned two Rottweilers, and Sue was afraid of dogs, so Sue never visited Tom's house."*, we do use taxonomic knowledge to recognise that Rottweiler is a type of dog; however, we also need detailed knowledge about ownership, domestic arrangements and human psychology that goes well beyond purely taxonomic information. And, to understand *"The boat sank because it was loaded with too much coal."* one needs knowledge of commonsense physics for which taxonomic relationships have little relevance.

Traditionally, knowledge needed to interpret such examples has been captured by augmenting the ontology with extensive axiom sets encoding commonsense background knowledge, as advocated by the *Naïve Physics Manifesto* [22] and exemplified by the CYC project [23]. An alternative, however, is to investigate how such knowledge might emerge from more fundamental conditions of reality, with higher-level concepts introduced incrementally by means of definitions grounded in those underlying structures. This latter approach may initially seem implausible, since high-level, human-oriented notions—such as being afraid of dogs—appear to be so far removed from fundamental concepts like space and time that it is difficult to conceive how they could be grounded in such primitives. And yet, a complete specification of the unfolding material states of the universe would surely determine not only all occurrences of boat-sinking events, but also all situations in which human fear of dogs is manifest—not only in actual encounters with dogs, but also in conversations, physiological responses such as muscular tension, or avoidance behaviours concerning dogs. The fact that such high-level phenomena are, in principle, fully determined by the underlying succession of physical states lends support to the view that they might ultimately be definable in terms of basic primitives.

The implausibility of actually formulating adequate definitions may be attributed to the apparent incompatibility between different levels of description—for example, between physical theories framed in terms of particles and forces, and the human-oriented perspective structured around physical objects and their affordances, or between material accounts of the world and those expressed in terms of beliefs, desires, and intentions. This apparent detachment of high-level, human-oriented concepts from material primitives—along with the presence of many evident semantic relationships at that level (such as "people avoid what they are afraid of")—has understandably led those constructing AI knowledge bases aimed at reproducing human reasoning to state axioms directly in terms of high-level concepts, rather than attempt the difficult task of grounding them in more fundamental terms.

In this paper, I take an alternative approach: rather than positing high-level concepts directly, I investigate how they might be built up in definitional layers from fundamental primitives. This strategy may be viable if we proceed incrementally and are able to bridge certain difficult conceptual gaps along the way. The focus will be on the foundations—specifically, how to introduce the notion of matter into a formal theory of space and time, and how this can serve as a basis for developing a theory of physical objects.

It should be noted that other approaches to grounding can be taken such as the *constructivism* of the Erlangen School [24, 25], which feed into more recent ideas for ontology developement and validation, such as proposed in [26]. These make different choices regarding what primitives should provide the grounding, which may be advantageous for formulating ontologies concerned with human-oriented conceptualisations.

## 2. Definitions and Definability

We now give an overview of *definitions* from a formal perspective and highlight some significant meta-logical properties and theorems. For current purposes, we confine attention to theories expressed in standard first-order logic. However, most of the notions introduced would apply to any of a broad class of logical languages.

### 2.1. The Definition of 'Definition'

**Definition 2.1.** *1st-order definitional formulae take one of the following forms, depending on the type of symbol defined:*

- $\forall x_1 \dots x_n[R(x_1, \dots, x_n) \leftrightarrow \Psi(x_1, \dots, x_n)]$
- $\forall x y_1 \dots y_n[(x = f(y_1, \dots, y_n)) \leftrightarrow \Psi(x, y_1, \dots, y_n)]$
- $\forall x[(x = c) \leftrightarrow \Psi(x)]$

*The* ***definienda*** *(terms defined) by such formulae are respectively: the relation $R$, the function $f$ and the constant $c$. The expression $\Psi(x_1, \dots, x_n)$ is called the* ***definiens****.*

*A* ***first-order definition*** *is any formula of one of the above forms in which the definiendum does not occur in the definiens.*

The notion of definability allows us to identify subsets of the vocabulary of a theory which can be taken as 'primitive' concepts:

**Definition 2.2.** *A* primitive vocabulary *for a theory $\Theta$ is a set $S \subseteq \text{Voc}(\Theta)$ such that all symbols in* $\text{Voc}(\Theta)$ *are definable within $\Theta$ from the symbols in $S$.*

**Definition 2.3.** *A* minimal primitive vocabulary *for a theory $\Theta$ is a primitive vocabulary for $\Theta$ of which no proper subset is also a primitive vocabulary for $\Theta$.*

### 2.2. Definability in Terms of Models — the Method of Padoa

Within the context of a formal theory it is natural to consider definitions and definability from a syntactic point of view. However, one can also gain considerable insight into these notions by considering how possible dentotations of concepts are related within the possible models of a theory.

*Padoa's Method* [27] provides a way to tell whether some term (i.e. a predicate or relation, symbol or a constant name) employed within in an axiomatic theory is *independent* of the others, or on the contrary is *definable* in terms of one or more of the remaining terms. In terms of the modern parlance of model theory, Padoa's key notion of independence can be specified as follows:

**Definition 2.4.** *Suppose we have an axiomatic theory $\Theta$ that makes use of a vocabulary of non-logical terms $V = \{t_1, t_2, \dots, t_n\}$, with $t \in V$ and $T \subseteq V \backslash \{t\}$. Then the term $t$ is* ***independent*** *from terms $T$ just in case we can find two models that satisfy all the axioms of $\Theta$ and agree exactly on the denotations of all the symbols in $T$ but disagree on the denotation of $t$. If there are no such models then $t$ is* ***dependent*** *on $T$, in the context of $\Theta$.*

It is readily apparent that $t$ being dependent on $T$ within $\Theta$ is a necessary condition for $t$ being definable from the terms $T$ in $\Theta$. If not then any proposed 'definition' of $t$ that fixed its interpretation in terms of the symbols $T$ would rule out some possible models of $\Theta$, and hence would not be purely definitional.

Whether the converse is true, that dependence is sufficient for definability, is not so clear. However, Tarski, in [28, 29] showed that dependence is both necessary and sufficient for definability in higher logic; and, subsequently, Beth [30] proved with his famous definability theorem that, if we are dealing with a first-order theory, we can be sure that any dependent symbol can be given a first-order definition:[3]

---

[3] Beth refers to Padoa's notion of *dependence* as *implicit definability*, and the existence of a first-order definition is called *explicit definability*. In this nomenclature, the theorem states that implicit and explicit definability coincide for first-order theories.

**Theorem 1.** *For every first-order theory $\Theta$, a symbol $\boldsymbol{t}$ is dependent on a set of symbols $T$ within $\Theta$ just in case $\Theta$ entails a first-order definition formula, whose definiendum is $t$ and with only non-logical symbols from $T$ occurring in its definiendum.*

### 2.3. Illustration of the use of Padoa's Method to Analyse Definability

We can illustrate this with a concrete example. Suppose we have the following set of human property and family relationship terms (where the notation $\mathsf{Pred}/n$ indicates that predicate Pred has arity $n$):

$$\{\mathsf{Male}/1,\ \mathsf{Female}/,\ \mathsf{IsParentOf}/2,\ \mathsf{IsMotherOf}/2,\ \mathsf{IsMarriedTo}/2,\ \mathsf{IsMotherInLawOf}/2\}$$

It should be clear that $\{\mathsf{Male}/1,\ \mathsf{Female}/,\ \mathsf{IsParentOf}/2,\ \mathsf{IsMarriedTo}/2\}$ provide a sufficient basis of terms to define the other two concepts. Moreover, if we had an axiom $\forall x[\mathsf{Male}(x) \leftrightarrow \neg\mathsf{Female}(x)]$, then we only need one of the two gender predicates. We also see that we could define $\mathsf{IsMotherOf}/2$ from $\mathsf{Female}/2$ and $\mathsf{IsParentOf}/2$ without needing the $\mathsf{IsMarriedTo}/2$ predicate.

We could show these dependencies either by considering which definitions would hold in a reasonable axiomatisation of these predicates or we can use Padoa's method. To use this method we just need to recognise that for any model — i.e. in this case a collection of humans for whom certain family relations hold. We can then see immediately that if we fix the only gender and parenthood relation then IsMotherOf will be fixed, but since IsMarriedTo is not fixed then IsMotherInLawOf is also not fixed. So the latter cannot be defined without using the former.

In many areas of mathematics and science, principles similar in form and purpose to Padoa's method are widely employed whenever one is interested in establishing what causal factors that are relevant to a particular effect. We may think that we have identified all factors relevant to that effect since, in cases that are equivalent in terms of these factors, the effect has been found to be the same. However, if we later find two cases that are equivalent with respect to those factors but differ in the effect observed, we must accept that there must be other relevant factors.

### 2.4. Categoricity, Monotransformability and Definability

The investigations of Tarski in his paper *Some methodological investigations on the definability of concepts* [28, 29] proved that the definability relationships among the terms of a theory is completely determined by certain properties concerning *isomorphisms* between models of that theory.

**Definition 2.5.** *An* ***isomorphism*** *between two models $\mathcal{M}$ and $\mathcal{N}$ (with the same signature) is a bijective mapping between their domains that preserves and reflects the interpretation of all non-logical symbols. That is, relations, functions, and constants are mapped identically under the correspondence.*

**Definition 2.6.** *A theory is* categorical *if and only if there is* ***at least one*** *isomorphism between any two models.*

**Definition 2.7.** *A theory is* monotransformable *if and only if given any two models, there is* ***at most one*** *isomorphism between them.*

Now we introduce a key property concerning the definitional power of theories:

**Definition 2.8.** *A theory $\Theta$ is* conceptually complete[4] *iff there is no categorical theory $\Theta'$ such that $\Theta \subset \Theta'$ and $\mathsf{Voc}(\Theta')$ contains at least one symbol that is not definable within $\Theta'$ in terms of the vocabulary $\mathsf{Voc}(\Theta)$.*

---

[4] Tarski uses the phrase 'complete with respect to its specific terms' rather than 'conceptually complete'.

So a ‘conceptually complete’ theory is one of which no *categorical* extension contains any new concept that is not definable (in the extended theory) from the vocabulary already present in the original theory. Note that if we dropped the restriction of considering only categorical extensions then no theory would be conceptually complete: it is always possible to add extra undefinable terms to a theory if we don’t care about exactly specifying their meaning.

Though conceptual completeness is concerned with definability in categorical extensions of a theory it is still a useful notion for theories that we do not intend to axiomatise up to categoricity. Given that any categorical axiomatisation of a new concept within a conceptually complete theory would entail a definition, it will normally be more sensible to define the concept rather than give a partial axiomatisation. Alternatively, if we do want to explicitly ensure categoricity with respect to an ontology (or perhaps to a subdomain of an ontology) we would need to use an ontology with a language richer than first-order logic, such as the General Formal Ontology (GFO) [31, 4].

It is worth noting that conceptual completeness of a theory does *not* mean that it is capable of defining all possible relationships over its domain. Indeed, if the domain is infinite there will be an uncountable number of possible predicate extensions, so these cannot all be defined within a language with a recursively enumerable set of well defined formulae. The property of conceptual completeness just means that the definitional power of the theory cannot be extended by adding more primitives.

## 2.5. Tarski’s Conceptual Completeness Theorem

Now we come to the key theorem proved in [28, 29] where Tarski showed that the model-theoretic property of monotransformability corresponds exactly to conceptual completeness:

**Theorem 2.** *A theory is conceptually complete if and only if it is monotransformable.*

This tells us that conceptual completeness amounts to the condition that no model of the theory incorporates multiple identical copies of itself. This is because if such copies existed, we could add concepts that would distinguish between copies.

A good example of a theory that is not monotransformable, but can be made so, is the theory of a strict discrete total order relation such that for every element there is both a greater and a smaller element. All models of such a theory will be isomorphic to the integers. However, we can map any model to itself in such a way that the order is preserved but each element gets mapped to a different element at some fixed shift up or down the order structure (e.g. we can map each element to its successor). But if we designate one element by a special name, say ‘0’, no such shift is possible without changing the denotation of ‘0’, so the theory becomes monotransformable. Hence, the theory is also conceptually complete. No concept can be added to the theory that is not already definable.

The following consequence of **Theorem 2** (also given in [28, 29]) is easy to comprehend, when we consider that once we add a fixed coordinate system to space, the only automorphism that preserves that coordinate system is the identity mapping:

**Theorem 3.** *Euclidean geometry with a fixed coordinate system stipulated by additional constants is monotransformable, and hence conceptually complete.*

It is likely that anyone with experience of constructing ontologies to support functionality of computational information systems, will consider that analysis of definability in terms of symmetries of categorical models can have no direct bearing on practical ontology design. They might say that ‘real’ ontologies cannot even avoid non-intended models, let alone ensure categoricity or monotransformability, so consideration of such abstruse meta-semantic properties cannot provide any useful guidance. I think this would be a misapprehension for the following reasons:

Firstly, the Method of Padoa can certainly be used in an informal and useful way. When considering vocabulary of related properties and relations it is surely useful to consider whether the interpretation of some subset of the vocabulary would fix the interpretation of the remaining terms. Then one could concentrate on axiomatising the subset (the primitives) while stipulating the meanings of the remaining terms by purely definitional axioms.

Secondly, although categoricity is not a property that one would expect in practical ontologies, this condition is not required for conceptual completeness. Monotransformability, is the key concept and implies that all terms would be definable *if* the theory were to be extended to a categorical theory, we do not actually need to do this. In fact for most ‘ordinary’ ontologies extending to a categorical theory is possible *in principle.* We would need to specify all the objects and the primitive (i.e. undefinable) properties and relations that hold. Although impractical, this must be *possible* for any consistent theory that can have a finite domain (e.g. of physical objects). It would also apply to theories that have domains, or sub-domains that are infinite but can be categorically axiomatised (e.g. a time series and/or spatial geometry).

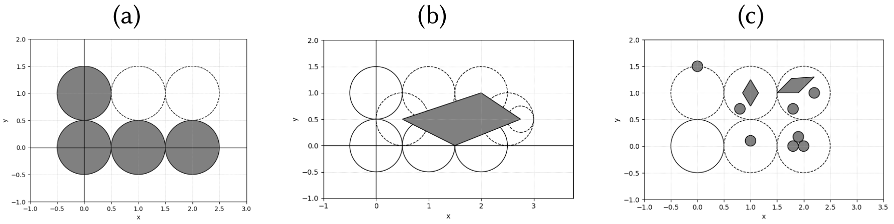

**Figure 1:** (a) coordinates anchored by discs, (b) a constructed shape, (c) a complex matter distribution.

## 3. Matter in Space

We cannot articulate our commonsense conceptualisation of the world in terms of vocabulary that describes only purely spatial and temporal aspects of reality. There needs to be some stuff in space to constitute physical objects; and the stuff needs to move around in order to manifest the processes and events that occur as the world evolves through time. Moreover, there does not seem to be any possibility that physical objects and their properties could be defined purely in terms of space and time, at least not in terms of conventional theories of space and time.

In the Newtonian theory, space and time characterise the structure of a void within which matter exists. They do not provide any properties that could determine the location of a particular kind of matter at a particular place in a particular time. And in terms of the Method of Padoa, we see that all purely spatial and temporal relationships could stay the same while the relationships pertaining to configurations of matter can vary. Of course matter is located in space, but space itself does not determine the location of matter.

### 3.1. Tarski’s Puzzle

The nature of foundational theories is that they describe the structure of reality in a way that is so general that they do not distinguish any specific features of the particular reality in which we live. Purely spatial, temporal and spatio-temporal relationships may apply to arbitrary regions of space (e.g. a region may be spherical, or overlap another region, or a spatio-temporal volume may correspond to a continuous path in space-time); but they do not pick out particular regions as being, say, occupied by a lump of stone, or corresponding to the trajectory of a football.

These considerations suggest that for a theory to support definitions of properties and relationships between physical objects it must include some primitives that are not purely spatio-temporal in character. But Tarski’s result regarding conceptual completeness of Euclidean geometry in [29] implies that any theory that includes Euclidean geometry can be formulated purely in terms of geometrical primitives. As Tarski notes in the conclusion to [29], this gives rise to a puzzle with regard to theories that describe the properties and behaviour of physical objects in space (e.g. Newtonian mechanics). On the one hand, it seems that we must need more than purely spatial concepts to capture matter and its mechanics; but on the other hand, the conceptual completeness of Euclidean geometry means that any further concepts added to the theory would in fact be definable from the original primitives.

Tarski suggested that the situation might indicate that formulation of an adequate physical theory requires a radically different kind of geometry. This may indeed be so. If instead we consider space and matter in the context of Einstein's General Relativity, we find a strong interdependence between 'space-time' and matter. Yet neither is fully dependent on the other and it is not possible to eliminate matter from its equations. We should note that there are more exotic theories (e.g. Loop Quantum Gravity), where matter does indeed arise from the structure of space time itself. However, these possibilities do not really explain the situation with regard to Euclidean geometry and classical mechanics. Surely it is possible to have a consistent classical mechanics incorporating standard geometry ([32] and [33] seem to demonstrate this). So the question of how to define matter within such a theory remains.

### 3.2. Solution to Tarski's Puzzle

I suggest that the solution to this puzzle is that the material content of the world can be introduced into the theory at the point where the coordinate system is fixed. As we saw above, achieving definitional completeness requires that a theory be made monotransformable. To do this we need to eliminate isomorphic copies within the structure, so that the only automorphism on the domain is the identity mapping. In the case of geometrical theories, this would normally be done by introducing some specific points or regions to act as an anchor for a coordinate system. But there is no reason why we should only introduce a few named spheres, as in Fig. 1(a). We can in fact introduce any designation whose structure is sufficiently complex to break all symmetries of empty space.

There are many ways in which this reference structure could be specified. But there is one way that enables us (despite **Theorem 3**) to enrich the theory with a concept of matter and extends its conceptual repertoire beyond that of purely spatial properties and relations. To do this we designate a particular sphere by a constant 'unit', and take its diameter as a unit of length; and, more significantly, we also introduce another constant 'matter', which has a dual purpose in the theory. One purpose is that 'matter' designates a region which we use to anchor the reference frame. But, as you may have guessed, we can also interpret 'matter' to be the extent of physical matter as it is distributed in space.

In order that the new primitives 'unit' and 'matter' fulfil their purposes we need to specify that they satisfy some axiom of the form:

$$\mathsf{Sphere}(\mathsf{unit}) \wedge \Phi(\mathsf{matter}, \mathsf{unit}) \ ,$$

where $\Phi$ describes a particular shape in space occupied by 'matter', whose size is calibrated by 'unit'.

In order for $\Phi$ to be suitable for picking out a particular spatial distribution, we need to ensure that $\Phi(r', \mathsf{unit})$ can only apply to regions of the same shape and size. Hence, $\Phi$ must satisfy the condition $\forall r'[\Phi(r', \mathsf{unit}) \rightarrow \mathrm{CG}(r', r)\ ]$, where CG is the relation of congruence (which in region-based Euclidean geometry is definable from parthood and sphericity primitives [34]). So, the complex predicate $\Phi(r, x)$ must use spatial primitives (e.g. parthood and sphericity) to specify a particular shape of spatial relation (e.g. it consists of spheres arranged in a particular way) and the size of this configuration is fixed in relation to that of the particular sphere designated by $x$. So when we name a particular sphere as unit and name a particular region $r$, which satisfies $\Phi(r, \mathsf{unit})$, as being the region matter, we fix a unique association between the vocabulary of the theory and particular spatial regions.

Note that the coordinate anchoring shown in Fig. 1(a) is a special case of this kind of construction, where 'unit' is the disc centred on the origin and 'matter' can be taken to be the sum of all the four discs. Moreover, we could take the shape constructed in Fig. 1(b) as the denotation of 'matter' and (again taking 'unit' as the disc at the origin) it would not be so difficult to specify a predicate $\Phi$ such that $\Phi(\mathsf{matter}, \mathsf{unit})$ would hold for this case, and $\Phi$ would also satisfy the given condition regarding congruence.

In view of the possibility of constructing a theory of space and matter of the form just described, we can understand the resolution of Tarski's puzzle as follows. We cannot fix space without introducing some reference points or regions that specify a coordinate system. Geometers have, as one would expect, solved this problem to their own satisfaction by specifying a coordinate system with purely geometrical elements (designated points or spheres to indicate a unit length and the orientation of axes).

But such a formulation of geometry cannot serve as a grounding for interpreting material properties of the world, since the reference markers are just arbitrary points or regions in space with no observable characteristics to distinguish them from any other points or regions in space. Basing the coordinate system on the location of matter in space makes sense, not only as a practical means of establishing a coordinate system but also as an explanation of how a formal theory of material objects and processes can be constructed in relation to a geometric theory that only describes empty space.

## 4. The Ontology of Matter

The observations of the previous section indicate that we can treat matter as a primitive notion independent from space; and, as well as providing a reference frame it can also provide a foundation for the definition of classes of physical objects and their properties and relations. Matter can either be considered as a continuous substance or as consisting of atoms. The decisive shift to atomism becoming the dominant view came only in the early 20th century after Einstein's 1905 paper on Brownian motion provided strong empirical evidence for atoms.

### 4.1. A Continuous Matter Theory

One way to ground a theory of physical reality and anchor it to space and time is simply to stipulate that there are one or more basic types of matter; and, at each point in time, each matter type occupies some particular region of space. This is essentially just an elaboration of the solution to Tarski's puzzle given above and illustrated in Fig. 1.

### 4.2. A Theory of Atoms and Molecules

Ontologies such as ChEBI [35] characterise chemical entities of a wide variety of types that are significant for chemists and particularly in relation to their role in biology. ChEBI also incorporates relationships of constituency that relate simple entities to complex ones, such as atoms to molecules. This is certainly very useful for a wide variety of applications where systematic organisation of information relating to chemistry and biology is important. However, ChEBI is largely Aristotelian in its content: it provides a classification of objects but (although it does indicate important constituency relationships) it does not explicitly define all its classes in terms of the structure and constituents of the members of those classes. Clearly doing that would be a huge and complex task. Nevertheless, there does not seem to be any reason why this would not be possible.

The following formulation of molecular structure and the class of molecules is only illustrative. It is strongly definitional in that it is constructed using formulae in which the only undefined primitive is the parthood relation, $\mathsf{P}(x, y)$, interpreted as 'Region $x$ is part of region $y$'. This presentation glosses over some significant complexities, in particular, the time dependence of certain predicates and relations.

- $\mathsf{O}(x,y) \ \equiv\ \exists m\,[\,\mathsf{P}(m,x) \wedge \mathsf{P}(m,y)\,]$
- $\forall xyz[\,(\mathsf{sum}(x,y) = z) \ \leftrightarrow\ \forall w\,[\,\mathsf{O}(w,z) \leftrightarrow (\,\mathsf{O}(w,x) \vee \mathsf{O}(w,y)\,)\,]$
- $\mathsf{At}(a) \ \equiv\ \forall x[\,\mathsf{P}(x,a) \ \rightarrow\ x = a\,]$

An object $x$ is a collection of objects of type $\kappa$ if and only if every part of $x$ overlaps with some object of type $\kappa$ which is part of $x$:

$$\kappa\_\mathsf{Coll}(x) \ \equiv_{def}\ \forall y[\mathsf{P}(y,x) \rightarrow \exists z[\mathsf{O}(y,z) \wedge \kappa(z) \wedge \mathsf{P}(z,x)]]$$

An aggregate $a$ of entities of a given type $\kappa$ with respect to a particular cohesion relation $\beta$ is a collection of entities of type $\kappa$, such that any two collections $a_1$ and $a_2$ whose sum is $a$ are such that we can find two entities of type $\kappa$ that are respectively parts of $a_1$ and $a_2$, and which are related by the cohesion relation.

$$\begin{aligned}\kappa\_\rho\_\mathsf{Agg}(x) \;\equiv_{def}\; & \kappa\_\mathsf{Coll}(x) \wedge \\ & \quad \forall x_1 x_2[\, (\kappa\_\mathsf{Coll}(x_1) \wedge \kappa\_\mathsf{Coll}(x_2) \wedge (x = \mathsf{sum}(x_1, x_2)) \rightarrow \\ & \qquad \exists y_1 y_2[\, \kappa(y_1) \wedge \kappa(y_2) \wedge \mathsf{P}(y_1, x_1) \wedge \mathsf{P}(y_2, x_2) \wedge \beta(y_1, y_2)]\end{aligned}$$

So a material structure constituted by a collection of atoms joined by atomic bonds can be characterised by the predicate At_AtBond_Agg. The cohesion constraint means that it is not possible to divide the collection into two sub-collections, such that there is no atomic bond linking the sub-collections.

$$\mathsf{Max}\_\varphi(x) \;\equiv_{def}\; \varphi(x) \wedge \forall y\, [\, (\varphi(y) \wedge \mathsf{P}(x, y)) \rightarrow x = y\, ]$$

So a molecule can be defined as a maximal aggregate of atoms that are joined by atomic bonds:

$$\mathsf{Mol}(x) \;\equiv_{def}\; \mathsf{Max\_At\_AtBond\_Agg}(x)$$

The atomic approach to matter can also be seen as a variant solution to Tarski's puzzle, where we fix space in relation to regions occupied by atomic or sub-atomic particles rather than types of matter distributed continuously over regions of space.

## 5. Conclusion

This paper has examined the definitional structure of formal ontologies with a particular focus on how material content can be incorporated into geometrical theories of space and time. Beginning from the model-theoretic foundations of definability, we reviewed key results including Padoa's method, Beth's theorem, and Tarski's analysis of conceptual completeness. These results provide not only tools for assessing the definitional coherence of a theory but also impose subtle constraints on the way new concepts—such as matter—can be introduced.

A central puzzle identified by Tarski is that Euclidean geometry is conceptually complete: any extension by new concepts must, if the theory is categorical, result in those concepts being definable from the original vocabulary. At the same time, it appears that purely geometrical notions are insufficient to characterise matter and physical objects. To address this apparent contradiction, we proposed that matter can be introduced at the point where a coordinate system is anchored—thus breaking model symmetries and fixing reference in a way that makes definitional completeness compatible with the inclusion of physical content.

Two approaches to the ontological representation of matter were considered. A continuous theory posits that matter types occupy regions of space at each time, allowing geometric anchoring through structured spatial distributions. Alternatively, a discrete, atomistic theory enables definitions of molecules and chemical structures as aggregates of atoms constrained by cohesion relations. Both approaches are compatible with the formal machinery of definitional ontology and suggest routes to reconciling classical physical theories with foundational logical frameworks.

This investigation has shown that the puzzle posed by Tarski can be resolved without abandoning classical geometry or ontology. Instead, by carefully considering the point at which semantic content is anchored, we can see how even seemingly qualitative features of the physical world—such as the presence of objects and substances—can be accounted for within a rigorously definitional framework. This opens the way for more grounded and formally precise constructions of mid-level ontologies that remain tied to fundamental physical structure.

Looking forward, two main directions suggest themselves for extending this approach. The first is to build upwards from the foundational treatment of matter by defining specific classes of physical objects in terms of structured configurations of material regions, and thereby incrementally extend conceptual terminology to incorporate mid-level vocabularies closer to those used in ordinary language. The second is to introduce one or more temporal primitives into the core foundation theory. These could then be used to define high-level concepts characterising events and processes.

While physical theories are typically axiomatized in terms of the evolution of matter over time, the perspective offered by our resolution of Tarski's puzzle suggests that physical laws themselves might be grounded in an extended matter-based coordinate system, one that encompasses not just spatial anchoring but also temporal structure. If we begin with a system that only fixes the geometry at a single time point, it is likely that one would again encounter the limitation of conceptual completeness, and be unable to enrich the theory with concepts that capture dynamic change. This suggests that the temporal structure of the world and the laws that govern it should be integrated from the outset into the axiomatic treatment of matter, so that the propensity for its distribution to evolve over time is built directly into its meaning.

## Declaration on Generative AI

During the preparation of this work, the author used ChatGPT both as an advanced search tool and to suggest or reword some sentences. The whole text has subsequently been thoroughly reviewed and edited by the author, who takes full responsibility for the publication's content.

## References

[1] R. Goclenius, Lexicon Philosophicum: Quo tanquam clave philosophiae fores aperiuntur, Typis Nicolai Hoffmanni, Frankfurt, 1613. Earliest known use of the term "ontologia" (ontology).

[2] R. Arp, B. Smith, A. D. Spear, Building ontologies with basic formal ontology, Mit Press, 2015.

[3] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, WonderWeb Deliverable D18: DOLCE Ontology Library, Technical Report, Laboratory For Applied Ontology - ISTC-CNR, 2003. http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf.

[4] H. Herre, General Formal Ontology (GFO): A foundational ontology for conceptual modelling, in: Theory and applications of ontology: computer applications, Springer, 2010, pp. 297–345.

[5] G. Guizzardi, G. Wagner, J. P. A. Almeida, R. S. Guizzardi, Towards ontological foundations for conceptual modeling: The unified foundational ontology (ufo) story, Applied ontology 10 (2015) 259–271.

[6] R. Carnap, Der logische Aufbau der Welt, Weltkreis, Berlin, 1928. Translated by R. George as *THe Logical Structure of the World*, University of California Press, 1969.

[7] B. A. W. Russell, The Problems of Philosophy, Williams & Norgate, London, 1912. Reprinted 2001, OUP.

[8] K. Fine, The study of ontology, Noûs 25 (1991) 263–294.

[9] S. Florio, O. Linnebo, Introduction to constructional ontology, in: I. J. da Silva Oliveira, P. P. F. Barcelos, R. F. Calhau, C. M. Fonseca, G. Righetti, G. Adamo (Eds.), Proceedings of the Joint Ontology Workshops (JOWO), FOUST: Workshop on Foundational Ontology, collocated with FOIS 2024, Enschede, The Netherlands, July 15-19, 2024, volume 3882 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024. URL: https://ceur-ws.org/Vol-3882/foust-5.pdf.

[10] B. Bennett, G. Sindoni, Ontology as structure, domain and definition, in: C. Trojahn, D. Porello, P. P. F. Barcelos (Eds.), Formal Ontology in Information Systems - Proceedings of the 14th International Conference (FOIS 2024), Enschede, The Netherlands, July 2024, volume 394 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2024, pp. 105–118. URL: https://doi.org/10.3233/FAIA241298. doi:`10.3233/FAIA241298`.

[11] E. Nagel, The Structure of Science: Problems in the Logic of Scientific Explanation, Harcourt, Brace & World, New York, 1961.

[12] J. A. Fodor, Special sciences (or: The disunity of science as a working hypothesis), Synthese 28 (1974) 97–115. doi:`10.1007/BF00485230`.

[13] J. Kim, Multiple realization and the metaphysics of reduction, Philosophy and Phenomenological Research 52 (1992) 1–26. doi:`10.2307/2108317`.

[14] M. A. Bedau, Weak emergence, in: J. E. Tomberlin (Ed.), Philosophical Perspectives: Mind, Causation, and World, volume 11, Blackwell, Oxford, 1997, pp. 375–399.

[15] T. O'Connor, H. Y. Wong, Emergence, in: E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy, 2002. URL: https://plato.stanford.edu/entries/emergence/, fall 2022 Edition.

[16] R. W. Batterman, Asymptotics and the role of minimal models, The British Journal for the Philosophy of Science 53 (2002) 21–38. doi:`10.1093/bjps/53.1.21`.

[17] M. Morrison, Emergent physics and micro-ontology, Philosophy of Science 79 (2012) 141–166. doi:`10.1086/663240`.

[18] N. Guarino, C. Welty, Evaluating ontological decisions with OntoClean, Commun. ACM 45 (2002) 61–65. doi:`http://doi.acm.org/10.1145/503124.503150`.

[19] R. Jackson, N. Matentzoglu, J. A. Overton, R. Vita, J. P. Balhoff, P. L. Buttigieg, S. Carbon, M. Courtot, A. D. Diehl, D. M. Dooley, W. D. Duncan, N. L. Harris, M. A. Haendel, S. E. Lewis, D. A. Natale, D. Osumi-Sutherland, A. Ruttenberg, L. M. Schriml, B. Smith, C. J. Stoeckert Jr., N. A. Vasilevsky, R. L. Walls, J. Zheng, C. J. Mungall, B. Peters, Obo foundry in 2021: operationalizing open data principles to evaluate ontologies, Database 2021 (2021) baab069. URL: https://doi.org/10.1093/database/baab069. doi:`10.1093/database/baab069`. `arXiv:https://academic.oup.com/database/article-pdf/doi/10.1093/database/baab069/408`

[20] I. Niles, A. Pease, Towards a standard upper ontology, in: Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001, 2001, pp. 2–9.

[21] M. Grüninger, Y. Ru, J. Thai, TUpper: A top level ontology within standards, Applied Ontology 17 (2022) 143–165.

[22] P. J. Hayes, The second naive physics manifesto, in: J. R. Hobbs, B. Moore (Eds.), Formal Theories of the Commonsense World, Ablex, 1985, pp. 1–36.

[23] D. B. Lenat, R. V. Guha, K. Pittman, D. Pratt, M. Shepherd, Cyc: Toward programs with common sense, Communications of the ACM 33 (1990) 30–49. doi:`10.1145/79173.79176`.

[24] R. Hesse, An introduction to the constructivist philosophy of the Erlangen School, International Philosophical Quarterly 19 (1979) 353–362. doi:`10.5840/ipq197919321`.

[25] P. Janich, Protophysics of time: Constructive foundation and history of time measurement, volume 30, Springer Science & Business Media, 2012.

[26] S. Scheider, J. A. Verstegen, What Is a Spatio-Temporal Model Good For?: Validity as a Function of Purpose and the Questions Answered by a Model, in: B. Adams, A. L. Griffin, S. Scheider, G. McKenzie (Eds.), 16th International Conference on Spatial Information Theory (COSIT 2024), volume 315 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2024, pp. 7:1–7:23. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.COSIT.2024.7. doi:`10.4230/LIPIcs.COSIT.2024.7`.

[27] A. Padoa, Un système nouveau des postulats fondamentaux de la théorie des nombres entiers, in: Comptes Rendus du Deuxième Congrès International des Mathématiciens, Paris, 1900, pp. 114–122. Presented in 1900, published 1902.

[28] A. Tarski, Uwagi metodologiczne o definowalności pojęć, Prace Towarzystwa Naukowego Warszawskiego, Wydział III Nauk Matematyczno-Fizycznych 34 (1931) 1–35. English translation in [29].

[29] A. Tarski, Some methodological investigations on the definability of concepts, in: Logic, Semantics, Metamathematics, Oxford Clarendon Press, 1956. Trans. J.H. Woodger.

[30] E. W. Beth, On Padoa's method in the theory of definition, Nederlandsche Akademie van Wetenschappen. Proceedings. Series A 56 (1953) 330–339. Also published in *Indagationes Mathematicae*, 15 (1953), 330–339.

[31] H. Herre, B. Heller, Semantic foundations of medical information systems based on top-level ontologies, Knowledge-Based Systems 19 (2006) 107–115.

[32] J. C. C. McKinsey, A. C. Sugar, P. C. Suppes, Axiomatic foundations of classical partical mechanics, Journal of Rational Mechanics and Analysis 2 (1953) 253–272.

[33] R. Montague, Deterministic theories, in: Decisions, values and groups 2, Pergamon Press, 1962. Reprinted in *Formal Philosophy: selected papers of Richard Montague*, ed. R. H. Thomason, Yale University Press, 1974.

[34] B. Bennett, A. G. Cohn, P. Torrini, S. M. Hazarika, A foundation for region-based qualitative geometry, in: W. Horn (Ed.), Proceedings of ECAI-2000, 2000, pp. 204–208.

[35] K. Degtyarenko, P. de Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. Alcántara,

M. Darsow, M. Guedj, M. Ashburner, Chebi: a database and ontology for chemical entities of biological interest, Nucleic Acids Research 36 (2008) D344–D350. URL: https://www.ebi.ac.uk/chebi/. doi:10.1093/nar/gkm791.