Adapting insect 'caste' concepts to the demands of bio-ontologies.

Thiago S. R. Silva* & Marcio R. Pie
Departamento de Zoologia, Universidade Federal do Paraná
*Corresponding Author's Email: tsranzanidasilva@gmail.com

Abstract

The present work provides a logical account for translating 'caste' concepts in social insects to an ontology-based data model, which can be used by researchers for describing and organizing entities belonging to non-human societies, as well as for the provision of evidential criteria for evaluating constitutive explanations of 'social' entities. We establish the top-level category for the concept of 'caste' and give examples on how to accommodate some subcategories (*e.g.* workers) in the ontology, following a domain granularity framework for the life sciences. We also provide accounts on current limitations in automated reasoning, current practices for 'caste' conceptualization, and improvements needed to be addressed in future works

Introduction

Terminological standardization has been a matter of debate among research groups that study insect societies for several years (Costa & Fitzgerald 1996, 2005; Peeters 2012; Neco *et* al. 2018; Silva & Feitosa 2019a; Sumner *et al.* 2018). According to several authors (*e.g.* Costa & Fitzgerald 2005; Crespi & Yanega 1995; Dew *et al.* 2016; Sumner *et al.* 2018), there are distinct reasons for the need for terminological univocity in this field of inquiry: (i) to provide a clearer conceptual landscape that enables researchers to explore drivers in the evolution of sociality in animals, (ii) to provide the linguistic components for a more straightforward communication, whether textually or orally, and (iii) to provide a framework for unified concept representation and organization, enhancing data comparability and reproducibility. Aiming for concept clarity is an important first step for standardization of specialized terminologies, although it does not directly entail standardization *per se*, since other terminological dimensions must be explored for consolidation of shared sets of terms (Epstein 2012; Faber 2015; Faber & León-Araúz 2016). While the above mentioned authors (and many others) provided distinct accounts on defining criteria for establishing unambiguous concepts of social categories in insect societies, few of them have tried to explore how these categories can be translated into data models. While urging for 'holistic' approaches for defining and describing social categories, researchers generally opt for combined frames of reference when defining 'castes', namely spatio-structural and functional frames (Michener 1974; Wilson 1975; Dew *et al.* 2016; Silva & Feitosa 2019a).

Although cross-referencing during concept representation can be considered the best approach for establishing conceptual boundaries of biological entities, authors normally equivocally ascribe concepts to certain frames in this process, leading to category mistakes (Ryle 2009). A category mistake is a semantic or ontological error made by an enunciator when it represents the facts of mental life as if they belonged to one logical entity or category (or domains of entities) when they actually belong to another (Ryle 2009; for examples of category mistakes in 'caste' representation see Figures 4 and 5 of Silva & Feitosa 2019a). According to Tanney (2009), researchers are partly led to construe mental concepts as signifying occurrences of underlying processes because they conflate how they explain an individual's successful moves with what they require of the individual in making those moves. In short, researchers make category mistakes because they normally conflate explanation with description. Despite being a simplified explanation on the causes of this semantic/ontological phenomena – since the origins of category mistakes are not so easily accounted for (Magidor 2019) – and considering that the process of description inevitably entails some sort of explanation, it is important to consider the implications of category mistakes in life sciences, especially if we intend to explore compositional approaches when representing biological concepts. Applying a logical framework that incorporates distinct frames of reference and levels of organization will help researchers account for the infelicity of category mistakes, at least to some degree.

Similar to other biological systems, insect societies have been investigated and understood within a general idea of levels of organization (Molet *et al* 2012; Strassman & Queller 2007; Sumner *et al.* 2018). The concept of levels can be broadly defined as the structure in which the natural world is perceived and organized, comprising several vertically stratified layers of entities and processes, such as the molecular, cellular, tissue, organ, organism, population, and ecosystem levels (Brooks *et al.* 2021). The general idea of levels of organization is extremely useful in several distinct contexts, ranging from descriptions to explanations and the provision of ontological inventories (List 2019), providing an important conceptual framework in various scientific and philosophical debates (Simon 1962; Schaffer 2003; Craver &

Bechtel 2007; Eronen 2013). In biology, various accounts of hierarchical compositionality of different levels of biological organization were proposed, in an attempt to answer how biological entities interacted with each other to form other entities belonging to more inclusive levels (Novikoff 1945; Wimsatt 1994; Heylighen 2000; Korn 2005).

With the growing need of researchers to manage large amounts of data with the help of computers and software applications, propositions of levels and hierarchies based on levels found their way to information science and ontology research (Vogt 2019). Accordingly, ontology researchers have developed their approaches to levels (*i.e.* granularity levels) and to different types of hierarchies based on levels (*i.e.* granular perspectives), while providing explicit criteria for identifying and demarcating different levels and different hierarchies (*i.e.* granularity framework) (Vogt 2019). An ontology consists of a set of terms with commonly accepted definitions that are formulated in a highly formalized canonical syntax and standardized format, yielding a lexical or taxonomical framework for knowledge representation (Smith 2003). They are interesting tools for representing and organizing specialized knowledge, especially when we are trying to arrange entities into a set of different levels of organization.

The terms in an ontology are organized into a nested hierarchy of classes and subclasses, forming a tree of increasingly specialized terms that is called a taxonomy (Rosse *et al.* 1998). However, when ontology researchers need to refer to hierarchies other than taxonomies, for example, a partonomy (i.e., a hierarchy based on part-whole relations), they usually do that in reference to some (external) granularity framework (Vogt 2019). Such partonomies, however, are usually only expressed indirectly through formalized descriptions specifying parthood relations between resources within the taxonomy of an ontology. This often results in the respective ontology containing several disconnected partonomies that provide only locally applicable parthood-based granularity schemes, as opposed to a single globally and universally applicable scheme (Vogt 2019).

The aim of this work is to provide a logical account for translating ‘caste’ concepts in social insects to an ontology-based data model, which can be used by researchers for describing and organizing entities belonging to non-human societies, as well as for the provision of evidential criteria for evaluating constitutive explanations of ‘social’ entities. We will adjust and refine previous propositions for a 'caste' ontology while addressing some misconceptions made by Silva & Feitosa (2019a). We will discuss the identity of the upper-level category of the 'caste' entity, based on the defining properties of some continuant entities (*i.e.* material entities and realizable entities) provided by the Basic Formal Ontology (BFO) and by Vogt *et al* (2012a). Additionally, we will explore 'caste' concepts in insect societies using distinct frames of reference and Vogt's (2019) domain granularity framework.

Methods

Specialized terms

Several terms routinely used in ontology development will be used recurrently throughout the text. We provide a brief explanation for most of them in Table 1; names available in the table will be underlined throughout the text. In this work, we will use quotation marks when referring to the name 'caste', in order to highlight the inappropriateness of its usage for communication in contemporary biology. We understand that reevaluating the linguistic component of terminological units[1] is essential for the development of a more inclusive and diverse science. We suggest that a reevaluation of this magnitude should be conducted by a diverse group of researchers through a transdisciplinary approach, which is a desirable condition needed to address linguistic modifications. Since it is not the aim of the present work to reevaluate the name usage throughout the history of the discipline, we refrain from providing an alternative term to the concept that refers to distinct categories of the phenotype in non-human societies, lest we change an oppressive-laden name by another.

Basic Formal Ontology for distinguishing material entities and realizable entities

The Basic Formal Ontology (BFO) serves as a template that defines types of entities and their divisions following a general granularity framework, providing the structure necessary for enabling cross-ontological comparability of application and domain reference ontologies (Vogt *et al.* 2012a). One of the main design principles for structuring BFO is the single inheritance model, which requires all defined categories to be disjoint and exhaustive (Vogt *et al.* 2012a), meaning that categories

---

[1] Here, we refer to the idea of multidimensionality of terms as understood in the Communicative Theory of Terminology (Cabré 2003), where a terminological unit is composed of three components: linguistic, cognitive and situational.

must be mutually exclusive relative to a given level of granularity (Spear 2006). This means that each class of a classification has maximally one single asserted parent class (Vogt *et al*. 2011).

The upper-level entities provided by the BFO classification can be distinguished into two types: continuant and occurrent entities. This distinction rests on a fundamental dichotomy between space and time (Smith 2015), being sufficient insofar for organizing the main axis of upper-level ontologies.

One of the main attributes of occurrent entities is that they can be bona fide or fiat in nature. They can be considered as a natural unit, however, only if they are parasitic on the existence of natural units in the continuant side (Smith 2015), and hence existing independently of human mental or linguistic activities (Vogt *et al.* 2012b). We can also identify sub-processes - *i.e.* temporal parts - which are fiat segments occupying constituent temporal intervals of the temporal interval occupied by the process as a whole (Smith 2015).

Most natural kinds are represented by continuant entities – such as organisms, parts of organisms, biological functions, roles – with material and realizable entities being subsets of a continuant. Contrary to occurrent entities, continuants can be delimited by bona fide boundaries depending on the perspective a certain entity is analyzed (Vogt *et al.* 2012b). This is mostly the case for material entities, which can be defined by properties belonging to several frames of reference rather than exclusively through a spatio-structural demarcation (Vogt *et al* 2012b).

As a subset of continuant entities, there are independent and specifically dependent entities (Figure 1A). The importance of the distinction between them is the way they are established through a relation of specific dependence. BFO establishes specific dependence as a relation that obtains between one entity and another when the first entity cannot exist unless the second entity exists also (Smith 2015). Hence, a continuant entity is dependent if, in order for it to exist, it must inhere in some other entity (Spear *et al* 2016). This relation can be one-sided or reciprocal.

In the BFO template, material entities subsume objects, fiat object parts, and object aggregates (Smith 2015), which assume a three-level theory of granularity. Since this template is considered inadequate for biology, Vogt *et al* (2012a) proposed additional types of material entities: object cluster, object group, fiat object part aggregate, fiat object part cluster, fiat object part group, object with fiat object part aggregate, object with fiat object part cluster, and object with fiat object part group.

Following the application of the single inheritance principle to the different sub-categories of a material entity, any given particular material entity must instantiate exactly one of the types of material entities defined for any given level of granularity (Figure 1B).

Dependent continuant entities are related to their bearers (*i.e.* independent continuant entities) by inherence, which in turn is defined as a one-sided existential dependence relation between two entities (Arp & Smith 2011). This means that realizable entities are only realized by some independent entity, at a particular time. The main subtypes of realizable entities are functions, roles, and dispositions (Figure 1C).

Vogt's domain granularity framework

Vogt's (2019) framework provides a basis for integrating distinct granular perspectives, through the identification of a set of causal unities that act as defining properties of a given entity. The main support of Vogt's proposition rests on the BFO's definition of 'object' (Smith *et al.* 2015) – a bona fide category that exists independent of human partitioning activities as causally relatively isolated entities (Ingarden 1983; Smith & Brogaard 2003) that are both structured through and maximal relative to a certain type of causal entity.

In the BFO, there are three distinct types of causal unities (*i.e.* causal unity via internal physical forces, causal unity via physical covering, and causal unity via engineered assembly of components) (Smith *et al.* 2015), with Vogt (2019) suggesting two types of causal unities that are suited to cover the missing cases for the life sciences (*i.e.* causal unity via bearing a specific function and causal unity via common historical/evolutionary origin).

In our case, four types of causal unities can be beneficial for establishing defining properties for 'caste' concepts: causal unity via internal forces, causal unity via physical covering, causal unity via bearing a specific function, and causal unity via common historical/evolutionary origin. Both causal unity via internal physical forces and causal unity via physical covering are associated with a spatio-structural granular perspective (Smith *et al.* 2015; Arp *et al.* 2015), while the causal unity via bearing a specific function is associated with a functional granular perspective and the causal unity via common historical/evolutionary origin is associated with a historical/evolutionary granular perspective (Vogt *et al.* 2012b).

In cases of cross-granular instantiation, in which the studied material entities do not necessarily directly sum to one another, integration of distinct granular perspectives can also be attained with Vogt's granularity framework, through several relations of granular representations.

Exploring these three granular perspectives and their underlying relations of causal unity for 'caste' conceptualization is beneficial (*e.g.* Silva & Feitosa 2019a), since the entities can be explored following distinct temporal partitions - retrodictive (diachronic) in an historical/evolutionary perspective, predictive in a functional perspective, and descriptive in a spatio-structural perspective (Figure 2).

Another important feature of Vogt's framework rests on the assumption that the entities that compose each level of a biological hierarchy can be represented as building blocks. These building blocks are representations of biological entities which are the sum of the building blocks belonging to finer levels of biological organization. Because the concept of a building block is based on an evolutionary interpretation, it explicitly predicts the diversification of newly evolved building blocks of a given level, with each higher level exhibiting the possibility of an exponentially larger number of different types of entities associated with a building block to be evolved—the number of possible types of molecules is exponentially larger than the number of possible types of atoms (Vogt 2019). When considering that actual material entities can be composed of several possible combinations (i.e. aggregates) of those building blocks, the diversity of possible types of material entities increases even more with each newly evolved building block.

Results

Result I: Top-level category of 'castes'

'Castes' are determined during some part of the development of organisms and maintained throughout their lives (Dolezal 2019; Trible & Kronauer 2017). The expression of a certain phenotypic trait is influenced by several processes acting in distinct levels of biological organization. Said traits can be determined equally or individually by epigenetic processes (Londe *et al* 2015), developmental and physiological determinants (Hartfelder *et al*. 2006), and/or environmental constraints (Peeters & Molet 2010). Hence, each component on a specific level of organization has a differential participation on the determination of particular 'castes'. Those components can (a) individually and directly influence the determination of the 'caste' entity, or (b) synergistically operate along with other components of the same level of organization they belong to, composing a coarser level entity that, in turn, will directly influence the development of the 'caste' entity. For example, one species of desert ant, *Cataglyphis mauritanica* (Emery, 1906a), has a strong genetic bias to 'caste' development, such that, under normal circumstances, certain genotypes always develop in small workers (Trible & Kronauer 2020). However, juvenile hormone treatment causes worker-destined genotypes to develop into larger queens, while queen-destined genotypes will develop into small workers when reared in small colonies, where the larvae are likely starved (Kuhn *et al*. 2018). Hence, the determination of 'castes' can be directed on the genotype level in some cases, while being directed at the metabolic or environmental level in other cases.

At this point, it is important to make a distinction between dispositions (*i.e.* biological functions) and processes, following BFO's template. Dispositions are exclusively described by reference to the types of process which would realize them under certain conditions (Ellis & Lierse 1994). Each disposition will exist in relation to some physical quality or qualities of its bearer; however, different quality patterns or arrangements may serve as ground under different circumstances or in different types of bearers (Spear *et al*. 2016). A biological function, as a subcategory of disposition, therefore, is realized only in reference to some type of process and performed by some material entity. Thus, one main attribute of a biological function is that, for a material entity to have the disposition to perform it does not necessarily imply that this entity is realizing this particular function at every moment in which it exists. Hence, biological functions are intimately related to the processes that realize them, but they are not identical or existentially dependent to them (Spear *et al.* 2016).

Processes and process boundaries, on the other hand, occupy spatiotemporal regions and they span temporal intervals and temporal instants, respectively (Smith 2012). Hence, processes are temporally extended, contrary to process boundaries. In the BFO template, processes and process boundaries (and their corresponding spatiotemporal regions and temporal regions) are considered as <u>occurrents</u> (Figure 3A). Processes are roughly defined as events that occur, unfold, happen, or develop through time, having temporal proper parts and are dependent of some continuant entity to happen (Jarrar & Ceusters 2017). Process boundaries, on the other hand, spans only zero-dimensional (*i.e.* temporal interval) and one-dimensional (*i.e.* temporal instants) temporal regions, meaning that they do not have temporal parts

(Jarrar & Ceusters 2017). An example of a biological process is the beating of a heart, while examples of process boundaries are the beginning and the ending of some organism's life. Temporal parts from a given process can be delimited by less inclusive process boundaries, such as the second and third years of an organism life. This type of temporal partition, however, can be problematic when we are trying to represent spatio-temporally extended entities, such as 'castes' (*cf.* Discussion I). Hence, any given material entity that is the bearer of a given function (in the form of having the disposition to perform it) participates in a given process, while the function is realized in the course of that same process (Figure 3B).

At some moment of their lives, particular 'castes' expressed in some individuals have the disposition to perform certain biological functions (Lillico-Ouachour & Abouheif 2017). These functions can be fixed for certain 'castes' (*e.g.* disposition for reproduction) or can be performed as a response to environmental constraints, such as local or temporal demands (*e.g.* food processing and defense behavior) (Shackleton *et al.* 2018; Klunk *et al.* 2020). Then again, 'castes' are expressed despite their dispositions to perform biological functions, which are considered sets of defining features for 'caste' categories rather than the 'castes' themselves.

Therefore, after determination, a 'caste' persists independently through time, maintaining their identity, without the necessity to establish a specific relation with another entity, determining it as independent continuant. Since 'castes' are studied and represented following distinct levels of material organization, they are categorized as material entities rather than realizable entities.

In short, 'castes' are phenotype categories with defining properties belonging to distinct frames of reference and granular perspectives, meaning that they are expressed during the development of an organism and maintained throughout their entire existence, being better understood and explored as material entities. To consider 'castes' as realizable entities (Silva & Feitosa 2019a) or as mechanisms (Sumner *et al.* 2018) denotes that they are expressed or organized only during some temporal partition of an organism lifespan and that this same organism has the disposition to belong to several 'castes' when certain demands compel them to, which, insofar, is not considered to be the case (*cf.* Discussion I).

Hence, following Vogt (2019) proposition for top-level categories of material entities, 'castes' can be considered as an 'epithelially delimited multi-cellular organism level entity' in a spatio-structural frame of reference, a 'functional unit' in a functional frame of reference, and an 'historical/evolutionary unit' in an historical/evolutionary frame of reference.

Result II: Frames and granular perspectives

Under the top-level categories of material entities (*i.e.* objects, fiat object part, object aggregate, fiat object part aggregate, and object with fiat object part aggregate; Vogt *et al.* 2012a), 'castes' are objects, and their bona fide boundaries can be established in a multicellular organism level. Hence, they are instantiated by fiat object part aggregates, objects with fiat object part aggregates, and objects at finer levels of representation – such as anatomical complexes, metabolic pathways, and genetic modularity. This way, the relation of each component of the phenotype to another, during the determination of 'castes', can be represented through a cumulative constitutive hierarchy (*i.e.* an hierarchical relation in which the parts of a multi-cellular organism that belongs to a cut of an instance granularity tree do not all instantiate the same basic type of phenotypic entity; Vogt 2019), creating several layers of 'caste' expression that do not necessarily directly sum to one another.

Vogt (2019) proposed five basic types of granular perspectives: a Compositional Building Block (CBB), Compositional Building Block Cluster (CBB-C), Region-based, Compositional Functional Unit (CFU), and Compositional Historical/Evolutionary Unit (CH/EU) granular perspectives. When modeling 'castes', at least three types of granular perspectives are important to us, namely CBB, CFU, and CH/EU. The other granular perspectives will probably be useful for dealing with some specific cases in which the concept is composed by aggregates of fiat entities from finer levels.

The CBB granular perspective encapsulates the main organizational axis of spatio-structural material entities, holding the most prototypical building blocks: 'atom' < 'molecule' (including metals and ionic compounds)< 'single-membrane-enclosed entity' (i.e. most organelles and all prokaryotic cells) < 'membrane-within-membrane entity' (i.e. eukaryotic cell) < 'epithelially-delimited compartment (i.e., some, but not all of the entities that are commonly referred to as organs) < 'epithelially-delimited multi-cellular organism' (i.e., organisms with an epidermis).

In our example, a spatio-structural worker is defined as an epithelially delimited multi-cellular organism composed by an epithelially-delimited compartment (in this case, an ovary). Hence, the ovary is the direct proper part of the spatio-structural worker (Figure 4A).

For a CFU granular perspective, we can consider five types of upper-level categories of functional entities: (i) functional unit of locomotion, (ii) functional unit of physiology, (iii) functional unit of ecology, (iv) functional unit of development, and (v) functional unit of reproduction and propagation. Three of them (categories i, ii, and iii) are related to dispositions independent of morphogenesis – namely, locomotory and physiological dispositions - and two of them (categories iv and v) are related to a morphogenetic disposition. Under the CFU, functional units are defined by their relation with several realizable entities. In our case, a functional unit related to a ‘caste’ (such as a worker) is defined by their disposition to perform certain roles or biological functions, following their upper level functional entities. As an example, a functional worker can be defined as a functional unit of ecology, since it has disposition to perform foraging behavior (Figure 4B).

For a CH/EU granular perspective, we can consider five types of upper-level categories of historical/evolutionary entities: (i) historical unit of development, (ii) historical unit of heredity, (iii) developmental lineage, (iv) genealogical lineage, and (v) evolutionary lineage. Two of them (categories i and ii) are related to structural integrity and stability over time – *i.e.* developmental and heredity relations –, while three of them (categories iii, iv, and v) are related to constituent historical relations distributed in time and space. Under the CH/EU, historical/evolutionary units are defined by their relation with chemical, biological, and/or historical entities. As an example, a historical/evolutionary unit related to a ‘caste’ is defined by several referential properties which are differentially expressed along a temporal continuum. As an example, a historical/evolutionary worker can be defined as an historical unit of development, since it can be developmentally induced by a juvenile hormone (Figure 4C).

One of the many limitations in the granularity schemes currently applied in bio-ontologies relates to the non-conformance with the reality of the biological organization of material entities (especially anatomical entities) (Vogt 2019). Since most granularity schemes applied in bio-ontologies presuppose an organization of material entities within a constitutive hierarchy, with each subcategory of a material entity standing in a direct subsumption relation to one another, the instance granularity tree will be directly translated into a type granularity tree. However, biological material entities are most possibly organized according to a cumulative constitutive hierarchy (Valentine & May 1996; Valentine 2004; Jagers Op Akkerhuis 2008).

When deriving granularity trees under a cumulative constitutive hierarchy, the mereological sum of all entities belonging to one instance granularity level does not necessarily sum to its unpartitioned whole, since the parts of a multi-cellular organism that belong to a cut of an instance granularity tree do not all instantiate the same basic type of anatomical entity (see Figure 2B, right, in Vogt 2019). Hence, the translation process of an instance granularity tree into a type granularity tree under a cumulative constitutive hierarchy is not necessarily straightforward. However, applying Vogt's (2019) sortation-by-type approach we can more easily derive type granularity trees of biological material entities.

For example, one can model several instances of a 'worker' entity in distinct granular perspectives through cross-granular representations (Figure 5). When integrating distinct representations of the same biological entity, in distinct frames of reference, we can sort them through relations of granular representation. If we are interested, for instance, in integrating the distinct subcategories of 'workers' represented in distinct frames of reference (spatio-structural, functional and historical/evolutionary), the spatio-structural representation can be logically integrated with both the functional representation and the historical/evolutionary representation, each belonging to their respective granular perspective (F-BR and H/E-BR), through relations of specific granular representations (has functional granular representation and has historical/evolutionary granular representation; Figure 5.A and Figure 5.B, respectively).

Building blocks that belong to finer levels in a CBB granular perspective can also be translated to other granular perspectives through granular relations. Hence, a gonad (or, more specifically in our case, an ovary) can be translated to a F-BR perspective through a granular relation of has functional granular representation, for example, being represented as a functional unit of reproduction and propagation in a functional frame of reference, having the disposition to perform some reproductive process.

Another important aspect of modeling biological data following granular perspectives is that any given biological material entity always instantiates several different material entity categories at the same time, one for each spatio-structural frame of reference (Vogt *et* al. 2012a). When building blocks are composed of aggregates or clusters of other entities, whether they possess bona fide or fiat boundaries in their respective granular perspective, they can be sorted through additional granular perspectives.

In the case of region-based granular perspectives (*i.e.* building block cluster, building block part, fiat building block aggregate, fiat building block part, fiat building block cluster, group of building block level objects, and fiat building block level entities), distinct perspectives can be sorted through proper parthood relations, sharing the same non-scale dependent single-relation-type (nrG) granularity type (Vogt 2019). According to Keet (2008), the nrG granularity type is a qualitative type of granularity that provides ordering of non-scale-dependent levels through a combination of properties where level identification is less straightforward. One distinguishing feature of the nrG granularity type is that it provides semantic aggregation (Keet 2008), defined as the combination of two or more semantic entities into one (Wilkinson 1995; Reape & Mellish 1999). Hence, when describing any set of region-based granular perspectives, the relations between the building blocks belonging to distinct granular perspectives can be logically and unambiguously sorted to a type granularity tree.

As two examples within our case study, we have the ‘female genitalia’ and the ‘gonadal tube’ of insect organisms in our granular perspective tree (Figure 5). The first occupy a region-based fiat building block aggregate granular perspective (Figure 5C), while the second occupy a region-based building block cluster granular perspective (Figure 5E), with each one being sorted to the corresponding level in a CBB granular perspective through their respective proper parthood relation (*i.e.* has proper part and proper part of).

Result III: Reasoning limitations

Although cross-granulation can be logically established following a domain granularity framework for life sciences, we still have some pressing issues relating to reasoning of data models (Mabee et al. 2020). To this date, ontology-based multi-species data models did not explicitly incorporated other frames of reference apart from a spatio-structural frame, providing only partial inventories of biological systems. However, with the growing need of logically modeling biological data in a cross-domain framework, aiming for a more throughout logical representation of biological systems, researchers found the necessity to explore reasoning ramifications of logical models that accommodated distinct frames within data models and data repositories (Mabee et al 2020).

Reasoning in ontology and knowledge bases roughly translates to the derivation of facts that are not explicitly expressed in said structures. Ontology reasoning is important because it enables designing and maintaining high quality models, enabling queries over ontology classes and their respective instances deposited at knowledge bases and makes the integration and alignment of multiple ontologies possible (Keet 2020). The process of automated reasoning (or simply reasoning) is made by softwares called reasoners.

Reasoners derive facts from different types of statements, usually from universal and/or particular statements. A set of universal statements comprise what is called a Terminological Box (TBox), while a set of particular statements is called an Assertion Box (ABox). In this sense, ontologies are composed of TBox expressions, while knowledge bases are composed of TBox and ABox expressions (De Giacomo & Lenzerini 1996). It is important that ontologies should be restricted to universal statements (TBox expressions), including terminological statements as metadata, with assertional and contingential statements (ABox expressions) being dealt in separate data repositories (Schulz & Jansen 2013), because the main objective of an ontology is to model statements that are universally true for all instances of a certain type of particular (Vogt & Bartolomaeus 2019).

According to Keet (2008), the current automated reasoners used in ontologies do not necessarily assume properly defined taxonomies of part-whole relations, since they only take into account the syntax of the relation-subrelation and do not consider the domain and range restrictions, nor the relational properties within the proposed taxonomy.

The most commonly used reasoner for bio-ontologies has been ELK, an OWL 2 EL compliant reasoner used to infer logical consequences in ontologies that have a large number of classes and/or properties. However, ELK follows a particular profile that is not suitable for modeling tasks in cross-domain ontologies, despite aiming for scalability. For example, the ELK reasoner does not support intransitive part-whole relations (*i.e.* despite x having a direct relation R to y and y having a direct relation R to z, x and z will not have a direct relation R to one another: $\forall(x,y: Rxy)\ \forall(z: Ryz)\ \neg Rxz$), which is a necessary conditions for partial ordering in data models that apply distinct granular perspectives.

One important issue that must be considered is that intransitive relations cannot be explicitly represented in OWL and the process of not asserting transitivity in this context means that a property is non-transitive (*i.e.* transitive in certain occasions and intransitive in others) and not necessarily intransitive (Keet 2014). It is important to notice that some types of biological entities can establish intransitive relations with other types of biological entities in specific circumstances (Guizzardi 2009; Vogt 2019) and ontologies that possess these types of relations are not necessarily ontologically flawed.

Hence, it is important that advances in reasoning architecture and inference rules are made in order to validate biologically coherent ontologies while maintaining ontological correctness. If we aim to model data following multiple frames of reference within a multi-domain ontology, we must explore alternatives for current reasoning practices. There will be always a trade-off, however, in expressiveness and applicability of logical reasoners that must be considered when modeling large-scale data structures (Mabee *et al*. 2020).

Discussion

Discussion I: Importance of defining 'castes' as material entities

Bio-ontologies are important and useful tools not only in the standardization of data and metadata, but also in data integration, data compatibility and comparability, and for data communication and management (Vogt *et al.* 2011). However, since several ontologies have been developed with a particular practical purpose in mind (*e.g.* Silva & Feitosa 2019a and this work), with a focus on the definition for very specialized types of entities, definitions for general types of entities are normally lacking. This normally leads to ontological inconsistency and cross-ontology incompatibility. Following an ontologically consistent framework for defining top-level categories in application oriented ontologies is important for enabling cross-ontological compatibility, also enhancing interoperability in future applications.

Another important point in defining 'castes' as top-level material entities is that, in this way, we can provide a conceptual framework for debates of downward causation on 'caste' expression (Craver & Bechtel 2007), since it is considered that all individuals in a colony share the same genetic background and that 'caste' differentiation derives from differences in gene expression (Crozier & Pamilo 1996; Elsner *et al.* 2018). Hence, reproductive disposition is considered to be decoupled from ‘caste’ expression in social insects. For example, in some groups of termites workers have the disposition to give birth in situations in which primary reproductives are absent (Korb & Hartfelder 2008; Korb 2015; Leniaud *et al* 2011). A similar case occurs in the ant genus *Dinoponera* Roger, 1861, although, in this case, the queen ‘caste’ is completely absent and reproductive workers (called *gamergates*) establish a relation of reproductive dominance among them (Peeters 1997). In these cases, historical/evolutionary (*e.g.* phenotype loss), ecological (*e.g.* relations of intracolonial dominance) and environmental (*e.g.* nest degradation or absence of food resources) factors may trigger reproductive disposition in workers.

Then again, due to the richness of functional dispositions of ‘caste’ entities, conceptualization under a functional frame of reference must be accounted for, when possible. As complex systems, social insects have the disposition to perform a series of intricate biological processes, which may vary and can be coupled to distinct ‘castes’. For example, Neves *et* al. (2017) observed that, in some species of ants (more precisely, in *Pheidole rudigenis* Emery, 1906b) there appears to be a higher variability in activity patterns due to the presence of more-than-one non-reproductive ‘caste’ (namely, soldiers/majors and workers/minors), while in species with only one non-reproductive ‘caste’ [*Gnamptogenys striatula* Mayr, 1884 and *Linepithema micans* (Forel, 1908)] activity patterns were less variable and more predictable. More-than-one non-reproductive ‘castes’ appears to be coupled, at least to some degree, to specific dispositions, such as performance of defensive or aggressive behaviors (Wu *et al*. 2018; Grüter *et al*. 2017). Although recognizing ‘castes’ by a set of functionally defined spatio-structural entities can be problematic (see discussion below), descriptive accounts of specific functionalities in certain ‘castes’ are extremely important for providing a robust framework for external explanatory accounts of complexity and, therefore, should be considered during modeling procedures.

There are, however, some limitations when describing ‘castes’ exclusively through functional or historical/evolutionary frames of reference. One prominent example refers to the concept of ‘temporal castes’ proposed by Wilson (1979). According to Wilson’s proposition, a ‘caste’ would be an ensemble of colony members that specialize on particular tasks for prolonged periods of time, being typically (but not necessarily) distinguished by other genetic, anatomical, or physiological traits. He further categorizes ‘castes’ in two subcategories, namely 'physical castes' and 'temporal castes', with the former being defined by allometric and other anatomic criteria and the latter being defined by age grouping. There are several ontological limitations in Wilson's proposition: (i) the top-level category of 'caste' used to subsume other finer categories of 'caste'; (ii) fortuitous categorical errors in Wilson's framework; (iii) misconceptions of the principle of persistence that allows the delineation of groups of temporal entities in Wilson's framework.

Relating to the first issue, Wilson assumes 'castes' as object aggregates (*i.e.* a group of bona fide epithelially delimited multicellular organisms) that are recognized by a set of functionally defined spatio-structural entities (*i.e.* anatomic distinct individuals that have the disposition to perform specific biological functions). However, recognition criteria of functionally defined spatio-structural entities cannot be directly inferred from their defining properties (*i.e.* dispositions), because said properties are not usually unambiguously bound to a specific set of spatio-structural properties (Vogt & Bartolomaeus 2019). According to Peeters (2012), certain groups of social insects possess this characteristic uncoupling of disposition (in his example, reproductive disposition) from a specific set of spatio-structural properties (*e.g.* development of morphoanatomical components associated with wings). Hence, recognizing 'castes' by a set of functionally defined spatio-structural entities constrains coherent accommodation of new findings relating to 'caste' expression and leads to the accretion of unnested categories in the classification (see further discussion below).

Wilson's second issue is, potentially, the most problematic one among the three of them, because it involves some misconceptions on the types of reasoning used in the process of proposing explanatory (retrodictive) accounts of evolutionary origin for 'caste' entities and the process of proposing predictive accounts of functional dispositions for 'caste' entities, while making several category mistakes when trying to translate these propositions of 'caste' expression to classificatory schemes. The author explores several underlying processes that explain how insect 'castes' are differentially expressed through the lifespan of a multicellular organism (*i.e.* age polyethism). In this process, he employs several types of biological entities, belonging to distinct frames of reference, in order to establish relations of causal unity between said entities that led to the expression and evolution of 'temporal castes'. When translating his explanations to a classificatory scheme, Wilson allocates entities belonging to one frame of reference to another, or allocates the same entity into two distinct frames. Superseding distinct frames of reference during description normally determines conceptual inconsistency, derived from category mistakes. Since each frame of reference virtually partitions the underlying biological entity in its own particular way, descriptions of the same phenotype that are based on different frames of reference often result in incongruent partitions (Vogt 2019). In our urge to provide a classificatory scheme that serve as a backbone to several explanatory claims, we normally mix up our descriptive demands with our explanatory demands.

Relating to Wilson's third issue, modeling concepts under a temporal continuum can be challenging and can inevitably lead us to proposing inconsistent categories. Different from spatio-structural entities, temporal entities cannot be directly and easily determined because their corresponding boundaries are not necessarily mind-independent, although they can be considered as bona fide entities in a given temporal perspective. Although discussions relating to temporal boundariness extends beyond the scope of this work, it is important to mention that, in order to recognize categories of temporal entities, one must assume that time has four basic distinct properties, according to Galton (2011), which it shares – to various degrees – with space: extension, linearity, directionality, and transience. An additional property of time relates to continuity, which presupposes extension, thus enabling the mapping of temporal-bounded biological entities to a spatio-structural frame of reference (although the mapping is not completely straightforward; *cf.* Vogt *et al.* 2012b for a throughout discussion on the matter). Each one of these properties presupposes another; extension presupposes linearity (*i.e.* time cannot be linear without being extended), linearity presupposes directionality (*i.e.* time cannot be directed without being linear), and directionality presupposes transience (*i.e.* time cannot be transient without being directed) (Galton 2011). In this sense, Wilson's proposition does not account for individual variation in a temporal continuum, but instead consider the temporal 'castes' (*i.e.* age cohort) as temporal scattered entities, in the form of fiat spatiotemporal regions, establishing some temporal instants that serve as partial fiat boundaries without taking into account the continuity property contemplated within their underlying temporal interval. Then, he fails to provide an account of connectedness that would enable the extension of said temporal scattered entities through time. Although temporal entities are inherently delimited by fiat boundaries to some extent, they can retain their bona fideness if their defining properties are properly accounted for. Wilson's third issue can be resolved if we represent temporal variation in the form of 'castes' as historical units of development within an historical/evolutionary frame of reference, instead of arbitrarily partitioning 'castes' in age cohorts.

Discussion II: Accommodating discoveries in the 'caste' data model and a brief discussion on 'subcastes'

Since the present data model is based on a domain granularity framework for the life sciences, all theoretically possible types of combinations of building blocks are covered (Vogt *et al.* 2011; Vogt 2019). Even if discoveries on 'caste' expression are made, this framework enables the reevaluation and evolution of concepts, attribution of new concepts and adjustments of available entities. Considering the major knowledge gap we have in understanding the drivers of 'caste' development and evolution, along with the huge advances we had in the past few years in fields like eco-evo-devo and thanks to the use of new high-throughput technologies, a versatile yet coherent framework helps us describe and accommodate new findings.

Another important benefit of using building blocks to model 'caste' concepts is its versatility when mapping entities to distinct frames of reference, without failing formal axioms. This is particularly important if we consider the wide range of possible dispositions and putative origins for each 'caste' entity and their underlying analytic value in machine-oriented applications. Each new finding can be accommodated to models if the structure of the ontology is based on these building blocks.

In groups that express intermediate or phenotypic mosaics, or have disposition to perform some specific functions or roles within a colonial environment (such as foragers, nurses, soldiers in bees, ants and termites), it is common to attribute a subcategory within the 'caste' concept, namely a 'subcaste'. In a tentative ontological framework, Silva & Feitosa (2019a) suggested that 'subcastes' can be defined as subcategories of 'caste' concepts. However, what is considered to be a 'caste' and a 'subcaste' sometimes depends on pragmatic reasons. If the 'subcaste' inherits all properties of its parent 'caste' and adds new defining properties and thus narrows in the meaning of the term, it truly represent a subclass of the former and, thus, could be argued to represent a 'subcaste' independent of our classification attempts.

Discussion III: Current limitations and further improvements

One limitation of the proposition for a ‘caste’ ontology relates to its specificity and small size. Since usage of ‘caste’ concepts is oriented to a small niche of researchers, an ontology will provide support only for particular set of demands. Nonetheless, due to an increase in complexity of the defining criteria needed to represent ‘castes’, establishing coherent sets of categories of ‘castes’ has been increasingly challenging. Alternatives to an ontology of ‘castes’ would be the description of individuals through knowledge bases using RDF specifications, contemplating several frames of reference, or the complete abandonment of the notion of ‘castes’ from insect societies. We believe, however, that the latter proposition is radical, to say the least; ‘caste’ concepts provide researchers with the much needed ontic support to provide explanatory accounts for spatio-structural, functional and historical/evolutionary variation and the complete abandonment of the idea would be uncalled for.

Further studies on category mistakes would be important to elucidate the causes of errors in attribution and their corresponding consequences in inferential processes in disciplines that use ‘caste’ concepts as analytical categories or as support assertions. Category mistakes can lead to problematic assertions, especially in situations that biological entities are understood and explored in more-than-one frame of reference; this can directly affect how underlying phenomena, such as functionality or evolutionary history, in biological entities are understood.

Another important limitation that needs to be stressed is the virtual absence of adequate reasoners to accurately infer logical consequences from intransitive axioms. Despite the expressiveness of OWL languages, especially in existential logics, they do not provide enough support for inferring logical consequences from cumulative-constitutive hierarchies. As discussed above, several biological entities are related in an intransitive manner, especially in particular granular perspectives, which are not contemplated in OWL formalisms. TBox reasoning does not have extensive or in-depth exploration when bio-ontologies are concerned; most works superficially mention which types of reasoners where used to validate ontology structure (Meehan *et al* 2011; Silva & Feitosa 2019b), while others briefly discuss some restrictions while evaluating bio-ontology consistency (Mungall *et al.* 2012), with only Dentler *et al*. (2011) providing a more in-depth survey and comparison of reasoners that succeed in classifying large biomedical ontologies. Normally, works have been made in reasoning over bio-ontologies, trying to improve reasoner efficiency while inferring logical consequences in knowledge bases (Balhoff *et al.* 2014, 2018; Mabee *et al.* 2020; Blondé *et al.* 2011). Future efforts in reasoners development should be made considering specific characteristics of biological entities.

Conclusion

Using a well-defined framework to organize ‘caste’ concepts enables coherent representation of several dimensions of biological reality. Although future efforts are needed to address specific issues in ‘caste’ modeling in the bio-ontology format, the establishment of biological aligned framework is a major step forward in dealing with much of the inconsistency pertaining ‘caste’ conceptualization.

Acknowledgements

The authors are deeply indebted to Alexandre Casadei Ferreira, Natalia Ladino, Mila F. O. Martins, Paloma L. Andrade, and Tainara T. Jory, who provided suggestions and discussions on the content of this manuscript. We are deeply grateful to Lars Vogt, who pointed out several inconsistencies in the 2019a work, leading to the development of this document. TSRS was funded by CAPES and Fundação Araucária (process number 88887.354448/2019-00).

References

Arp, R., Smith, B. (2011) Realizable entities in basic formal ontology. In: Nature proceedings, http://ontology.buffalo.edu/smith/articles/realizables.pdf

Arp R, Smith B, Spear AD. (2015) Building ontologies with basic formal ontology. Cambridge, Massachusetts: The MIT Press. p. 248.

Balhoff JP, Dececchi TA, Mabee PM, Lapp H. (2014). Presence-absence reasoning for evolutionary phenotypes. arXiv preprint arXiv:1410.3862. 2014 Oct 14.

Balhoff, J. P., Good, B., Carbon, S., & Mungall, C. (2018). Arachne: an OWL RL Reasoner Applied to Gene Ontology Causal Activity Models (and Beyond). In International Semantic Web Conference (P&D/Industry/BlueSky).

Blondé, W., Mironov, V., Venkatesan, A., Antezana, E., De Baets, B., & Kuiper, M. (2011). Reasoning with bio-ontologies: using relational closure rules to enable practical querying. Bioinformatics, 27(11), 1562-1568.

Brooks, D. S., DiFrisco, J., & Wimsatt, W. C. (2021). Levels of organization: the architecture of the scientific image. *In*: Brooks, D. S.. DiFrisco, J. & Wimsatt, W.C. *Levels of Organization in the Biological Sciences*. MIT Press, Cambridge.

Cabré, M. T. (2003). Theories of terminology: Their description, prescription and explanation. Terminology, 9(2), 163-199.

Costa, J. T., & Fitzgerald, T. D. (1996). Developments in social terminology: semantic battles in a conceptual war. Trends in Ecology & Evolution, 11(7), 285-289.

Costa, J. T., & Fitzgerald, T. D. (2005). Social terminology revisited: Where are we ten years later?. In Annales Zoologici Fennici (Vol. 42, No. 6, pp. 559-564). Finnish Zoological and Botanical Publishing Board.

Craver CF, Bechtel W. (2007) Top-down causation without top-down causes. Biol Philos., 22:547–63.

Crespi, B. J., & Yanega, D. (1995). The definition of eusociality. Behavioral Ecology, 6(1), 109-115.

Crozier, R.H., and Pamilo, P. (1996). Evolution of Social Insect Colonies: Sex Allocation and Kin Selection (Oxford: Oxford University Press).

De Giacomo, G., & Lenzerini, M. (1996). TBox and ABox reasoning in expressive description logics. KR, 96(316-327), 10.

Dentler K, Cornet R, Ten Teije A, De Keizer N. (2011) Comparison of reasoners for large ontologies in the OWL 2 EL profile. Semantic Web. 2011 Jan 1;2(2):71-87.

Dew, R. M., Tierney, S. M., & Schwarz, M. P. (2016). Social evolution and casteless societies: needs for new terminology and a new evolutionary focus. Insectes Sociaux, 63(1), 5-14.

Dolezal, A. G. (2019). Caste determination in arthropods. In: Encyclopedia of Animal Behavior (pp. 691-698). Elsevier.

Ellis, B. & Liese C. (1994). Dispositional essentialism. Australasian Journal of Philosophy, 72(1), 27–45.

Elsner D, Meusemann K, Korb J. (2018) Longevity and transposon defense, the case of termite reproductives. Proc Natl Acad Sci U S A., 115:5504–9.

Emery, C. (1906a). Rassegna critica delle specie paleartiche del genere Myrmecocystus. Memorie della Reale Accadémia delle Scienze dell'Istituto di Bologna (6)3:47-61 [pagination of separate: 173-187].

Emery, C. (1906b). Studi sulle formiche della fauna neotropica. XXVI. Bullettino della Società Entomologica Italiana 37:107-194.

Epstein, M. (2012). Nomenclature, terminology and language. Bionomina, 5(1), 1-56.

Eronen MI. (2013) No levels, no problems: downward causation in neuroscience. Philos Sci., 80:1042–52.

Faber, P. (2015). Frames as a framework for terminology. Handbook of terminology, 1(14), 14-33.

Faber, P. & León-Araúz, P. (2016) Bridging the Gap between Description and Standardization: a Frame-based Version of NATO Glossaries. In Proceedings of the XVII EURALEX International Congress. Lexicography and Linguistic Diversity, edited by Margalitadze, T. & Meladze, G., pages 499-508. Tbilisi: Ivane Javakhishvili Tbilisi University Press.

Forel, A. (1908). Ameisen aus Sao Paulo (Brasilien), Paraguay etc. gesammelt von Prof. Herm. v. Ihering, Dr. Lutz, Dr. Fiebrig, etc. Verhandlungen der Kaiserlich-Königlichen Zoologisch-Botanischen Gesellschaft in Wien 58:340-418.

Galton, A. (2011). Time flies but space does not: Limits to the spatialisation of time. Journal of Pragmatics, 43, 695–703.

Grüter, C., Segers, F. H., Menezes, C., Vollet-Neto, A., Falcón, T., von Zuben, L., ... & Almeida, E. A. (2017). Repeated evolution of soldier sub-castes suggests parasitism drives social complexity in stingless bees. Nature communications, 8(1), 1-8.

Guizzardi, G. (2009). The problem of transitivity of part-whole relations in conceptual modeling revisited. In International Conference on Advanced Information Systems Engineering (pp. 94-109). Springer, Berlin, Heidelberg.

Hartfelder, K., Makert, G. R., Judice, C. C., Pereira, G. A., Santana, W. C., Dallacqua, R., & Bitondi, M. M. (2006). Physiological and genetic mechanisms underlying caste development, reproduction and division of labor in stingless bees. Apidologie, 37(2), 144-163.

Heylighen, F. (2000). Evolutionary Transitions: how do levels of complexity emerge?. COMPLEXITY-NEW YORK-, 6(1), 53-57.

Ingarden R. (1983) Man and value. München: Philosophia Verlag GmbH. p. 184.

Jagers Op Akkerhuis, G.A.J.M. (2008) Analysing hierarchy in the organization of biological and physical systems. Biological Reviews of the Cambridge Philosophical Society 83(1): 1–12.

Jarrar, M, & Ceusters, W. (2017) Classifying processes and basic formal ontology. In: ICBO 2017. Available at: http://ceur-ws.org/Vol-2137/paper_14.pdf

Keet CM (2008) A formal theory of granularity - toward enhancing biological and applied life sciences information system with granularity. Bozen: Free University of Bozen - Bolzano. Available at http://www.meteck.org/files/AFormalTheoryOfGranularity_CMK08.pdf.

Keet, C. M. (2014). Preventing, detecting, and revising flaws in object property expressions. Journal on Data Semantics, 3(3), 189-206.

Keet, C.M. 2020. An introduction to ontology engineering. [Online]. Version 1.5. [Accessed 13 April 2021]. Available from: https://people.cs.uct.ac.za/~mkeet/files/OEbook.pdf

Klunk C.L., Argenta M.A., Casadei-Ferreira A., Economo E.P., Pie M.R. (2020) Mandibular morphology, task specialization, and bite mechanics in Pheidole ants (Hymenoptera: Formicidae) bioRxiv. 11.16.385393 doi: https://doi.org/10.1101/2020.11.16.385393

Korb J. (2015) A central regulator of termite caste polyphenism. Adv Insect Physiol. 83:295–313.

Korb J & Hartfelder K. (2008) Life history and development-a framework for understanding developmental plasticity in lower termites. Biol Rev. 83:295–313.

Korn, R. W. (2005). The emergence principle in biological hierarchies. Biology and Philosophy, 20(1), 137-151.

Kuhn, A., Darras, H., & Aron, S. (2018). Phenotypic plasticity in an ant with strong caste–genotype association. Biology letters, 14(1), 20170705.

Leniaud L, Darrouzet E, Dedeine F, Ahn K, Huang Z, Bagneres AG. (2011) Ontogenic potentialities of the worker caste in two sympatric subterranean termites in France. Evol Dev. 13:138–48

Lillico-Ouachour, A., & Abouheif, E. (2017). Regulation, development, and evolution of caste ratios in the hyperdiverse ant genus Pheidole. Current opinion in insect science, 19, 43-51.

List, C. (2019). Levels: descriptive, explanatory, and ontological. Noûs, 53(4), 852-883.

Londe, S., Monnin, T., Cornette, R., Debat, V., Fisher, B. L., & Molet, M. (2015). Phenotypic plasticity and modularity allow for the production of novel mosaic phenotypes in ants. Evodevo, 6(1), 1-15.

Mabee PM, Balhoff JP, Dahdul WM, Lapp H, Mungall CJ, Vision TJ. A logical model of homology for comparative biology. Systematic biology. 2020 Mar 1;69(2):345-62.

Magidor, O (2019) Category Mistakes, The Stanford Encyclopedia of Philosophy (Winter 2020 Edition). Acessado em: 28/04/2021. Disponível em: https://plato.stanford.edu/archives/win2020/entries/category-mistakes.

Mayr, G. (1884). [Untitled. Descriptions of eight new species.]. Pp. 31-38 in: Radoszkowsky, O. 1884. Fourmis de Cayenne Française. Trudy Russkago Entomologicheskago Obshchestva 18:30-39.

Meehan, T. F., Masci, A. M., Abdulla, A., Cowell, L. G., Blake, J. A., Mungall, C. J., & Diehl, A. D. (2011). Logical development of the cell ontology. BMC bioinformatics, 12(1), 1-12.

Michener CD (1974) The social behavior of the bees. Harvard University Press, Cambridge

Molet, M., Wheeler, D. & C. Peeters (2012). Evolution of novel mosaic castes in ants: Modularity, phenotypic plasticity, and colonial buffering. The American Naturalist 180: 328-341.

Mungall CJ, Torniai C, Gkoutos GV, Lewis SE, Haendel MA. (2012) Uberon, an integrative multi-species anatomy ontology. Genome biology. 13(1):1-20.

Neco, L. C., Japyassú, H. F., El-Hani, C. N., & Châline, N. (2018). From classificatory to quantitative concepts in the study of sociality in animals: an epistemological view. Biological Theory, 13(3), 180-189.

Neves, F. M., Viana, R. L., & Pie, M. R. (2017). Recurrence analysis of ant activity patterns. PloS one, 12(10), e0185968.

Novikoff, A. B. (1945). The concept of integrative levels and biology. Science, 101(2618), 209-215.

Peeters C. (1997). Morphologically primitive ants: comparative review of social characters, and the importance of queen-worker dimorphism. In: Choe JC, Crespi BJ, eds. The evolution of social behavior in insects and arachnids. Cambridge: Cambridge University Press, 372–391.

Peeters, C. (2012). Convergent evolution of wingless reproductives across all subfamilies of ants, and sporadic loss of winged queens (Hymenoptera: Formicidae). Myrmecological News, 16, 75-91.

Peeters, C. & Molet, M. 2010. Colonial reproduction and life histories. pp. 159-176 In: Ant Ecology (eds L. Lach, C. Parr and K. Abbott). Oxford University Press.

Reape, M. & Mellish, C. (1999) Just what is aggregation anyway? In: Proceedings of the European Workshop on Natural Language Generation (EWNLG 1999), Toulouse, France, pp. 20–29.

Roger, J. (1861). Die Ponera-artigen Ameisen (Schluss). Berliner Entomologische Zeitschrift 5:1-54.

Rosse C, Mejino JL, Modayur BR, Jakobovits R, Hinshaw KP, et al. (1998). Motivation and organizational principles for anatomical knowledge representation: the digital anatomist symbolic knowledge base. J Am Med Inform Assoc. 5:17–40.

Ryle, G. (2009) The Concept of Mind. 60th Anniversary Edition. London: Routledge.

Schaffer J. (2003) Is there a fundamental level? Noûs, 37:498–517

Schulz, S. & Jansen, L. (2013) “Formal ontologies in biomedical knowledge representation.,” IMIA Yearbook of medical informatics 2013, vol. 8, no. 1, pp. 132–46.

Shackleton, K., Alves, D. A., & Ratnieks, F. L. (2018). Organization enhances collective vigilance in the hovering guards of *Tetragonisca angustula* bees. Behavioral Ecology, 29(5), 1105-1112.

Silva, T.S.R. & Feitosa R.M. (2019a). On titles and royalty: a terminological discussion over castes in myrmecology. Insectes sociaux, 66(1): 25-35.

Silva, T.S.R. & Feitosa, R. M. (2019b). Using controlled vocabularies in anatomical terminology: a case study with *Strumigenys* (Hymenoptera: Formicidae). Arthropod structure & development, 52, 100877.

Simon H. (1962) The architecture of complexity. Proc Am Philos Soc.106:467–82.

Smith B. (2003) Ontology. In: Floridi L, editor. Blackwell guide to the philosophy of computing and information. Oxford: Blackwell Publishing. p. 155–66.

Smith B. (2012) Classifying Processes: An Essay in Applied Ontology. Ratio, 25:4, 463-488

Smith B, Brogaard B. (2003) Sixteen days. J Med Philos., 28:45–78.

Smith, B. (2015). Basic Formal Ontology 2.0: Specification and user’s guide. Available at: https://github.com/bfo-ontology/BFO/wiki. Accessed at: 18.v.2021.

Smith B, Almeida M, Bona J, Brochhausen M, Ceusters W, et al. (2015) Basic formal ontology 2.0. Available: https://github.com/BFO-ontology/BFO2.

Spear AD (2006) Ontology for the Twenty First Century: An Introduction with

Recommendations. Science. pp 1–132.

Spear, A. D., Ceusters, W., & Smith, B. (2016). Functions in basic formal ontology. Applied Ontology, 11(2), 103-128.

Sumner, S., Bell, E., & Taylor, D. (2018). A molecular concept of caste in insect societies. Current opinion in insect science, 25, 42-50.

Strassmann, J. E. & Queller, D. C. (2007). Insect societies as divided organisms: the complexities of purpose and cross-purpose. Proceedings of the national academy of sciences, 104(suppl 1), 8619-8626.

Tanney, J. (2009) Rethinking Ryle: A Critical Discussion of the Concept of Mind. In: Ryle, G. The Concept of Mind. 60th Anniversary Edition. London: Routledge.

Trible, W., Kronauer, D.J.C. (2017). Caste development and evolution in ants: it's all about size. Journal of Experimental Biology 220, 53-62

Trible, W. & Kronauer, D. J. (2020). Hourglass Model for Developmental Evolution of Ant Castes. Trends in ecology & evolution, 36 (2): 100-103.

Valentine J.W. (2004) On the origin of phyla. Chicago: University of Chicago Press. pp 381–424.

Valentine, J. W., & May, C. L. (1996). Hierarchies in biology and paleontology. Paleobiology, 23-33.

Vogt, L. (2019). Levels and building blocks—toward a domain granularity framework for the life sciences. Journal of Biomedical Semantics, 10(1), 4.

Vogt, L. & Bartolomaeus, T. (2019). Adapting Ontologies to the Diagnostic Needs of Anatomy: Revisiting the Type Concept in Biology. Available at: https://doi.org/10.13140/RG.2.2.14892.51844

Vogt, L., Grobe, P., Quast, B., & Bartolomaeus, T. (2011). Top-level categories of constitutively organized material entities-suggestions for a formal top-level ontology. PLoS One, 6(4), e18794.

Vogt, L., Grobe, P., Quast, B., & Bartolomaeus, T. (2012a). Accommodating ontologies to biological reality—top-level categories of cumulative-constitutively organized material entities. PLoS One, 7(1), e30004.

Vogt, L., Grobe, P., Quast, B., & Bartolomaeus, T. (2012b). Fiat or bona fide boundary—a matter of granular perspective. PLoS One, 7(12), e48603.

Wilkinson, J. (1995) Aggregation in Natural Language Generation: Another Look. Technical report, Computer Science Department, University of Waterloo.

Wilson, E.O. (1975). Sociobiology: The New Synthesis. Harvard University Press.

Wilson, E. O. (1979). The evolution of caste systems in social insects. Proceedings of the American Philosophical Society, 123(4), 204-210.

Wimsatt, W. C. (1994). The ontology of complex systems: levels of organization, perspectives, and causal thickets. Canadian Journal of Philosophy, 24(sup1), 207-274.

Wu, T, Dhami, GK, Thompson, GJ. (2018) Soldier-biased gene expression in a subterranean termite implies functional specialization of the defensive caste. Evolution & Development. 20: 3-16.

Tables

Table 1. Glossary of terms related to ontology development and computational reasoning.

Name	**Explanation**
Bona fide boundary	Natural or mind-independent boundaries, which are physical boundaries in the things themselves that exist independently from human perception.
Building block	A 'Lego-brick-like' entity that evolves, diversifies, and provides reality's inventories of basic categories of material entities. They are spatio-structurally, functionally, developmentally, and evolutionarily both integrated and stable, while increasing nature's overall evolvability. They can interact with other building blocks to form aggregates and more complex building blocks, especially when analyzed in different granular perspectives.
Continuant entity	An entity that persists, endures, or continues to exist through time while maintaining its identity. These entities incorporate both material and immaterial continuants extended and potentially moving in space, and the spatial regions at which they are located and through which they move, and their associated spatial boundaries.
Disjointness	In an ontology, disjointness relates to two categories belonging at the same level of granularity that cannot share an instance.
Exhaustiveness	Completeness of a given ontology in terms of types of entities and types of relations by which entities are tied together to form large wholes.
Fiat boundary	Artificial (*i.e.* artifact of cognition) or mind-dependent boundaries, which are non-physical boundaries that depend on human decision and thus are the products of mental activities.
Frame of reference	General definition: a unit or organization of units that serve to identify a coordinate system with respect to which certain properties of objects, including the phenomenal self, they are gauged. This definition is modified depending on the domain of inquiry in which

	frames are used. Ontology development definition: a data structure that contains all the information in the ontology about a given domain of inquiry. It denotes a set of representative resources that provides a baseline value against which an ontology should be compared.
Granularity framework	A static structure characterized by a set of granular levels and hierarchies (*i.e.* granular perspectives) in a given subject domain.
Granularity levels	A set of several vertically stratified layers of entities and processes, accommodated within a given granular perspective.
Granular perspective	An hierarchy containing a set of vertically stratified layers of entities and processes.
Material entity	A continuant entity that is spatially extended and whose identity is independent of that of other entities and can be maintained through time.
Occurrent entity	An entity that unfolds itself in time, or it is the instantaneous boundary of such an entity (*e.g.* a beginning or an ending), or it is a temporal or spatiotemporal region. Occurrent entities relate to processes, boundaries of processes, or spatial-temporal regions and can be arbitrarily summed and divided.
Realizable entity	A specifically dependent continuant that inheres in other continuant entities and is not exhibited in full at every time.
Reasoner	A piece of software able to infer logical consequences from a set of asserted facts or axioms.
Single inheritance	An ontological principle that establishes that every universal or class included in a given classification should stand in an inheritance relationship to exactly one universal or class at the next highest level.

Figures

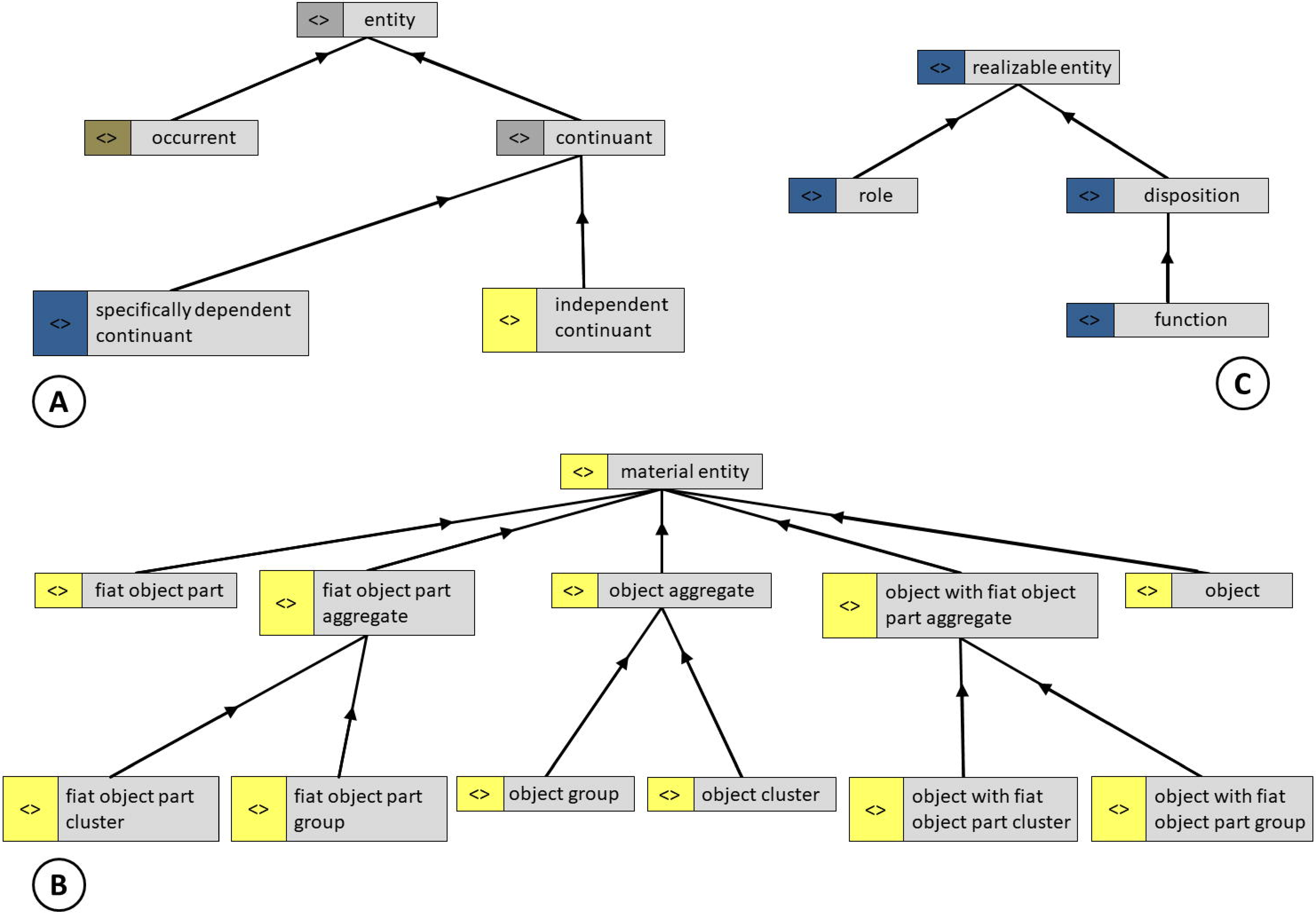
Fig 1 Distinct top-level categories of entities and their respective relations following BFO’s template and Vogt *et al.* (2012a) proposition. Top-level categories of entities: A. Top-level categories of entities, representing occurrent, continuant, independent, and specifically dependent entities and their respective relations. B. Top-level categories of material entities, showing the relations among BFO's and Vogt *et al* (2012a) material entities. C. Top-level categories of realizable entities, showing the relations among BFO's realizable entities. Arrows indicate an *is_a* relation property. Colors in the ‘*more-than less-than*’ boxes indicate the types of the top-level categories each entity is subsumed into: yellow are independent continuants, blue are specifically dependent continuants, and brown are occurrents.

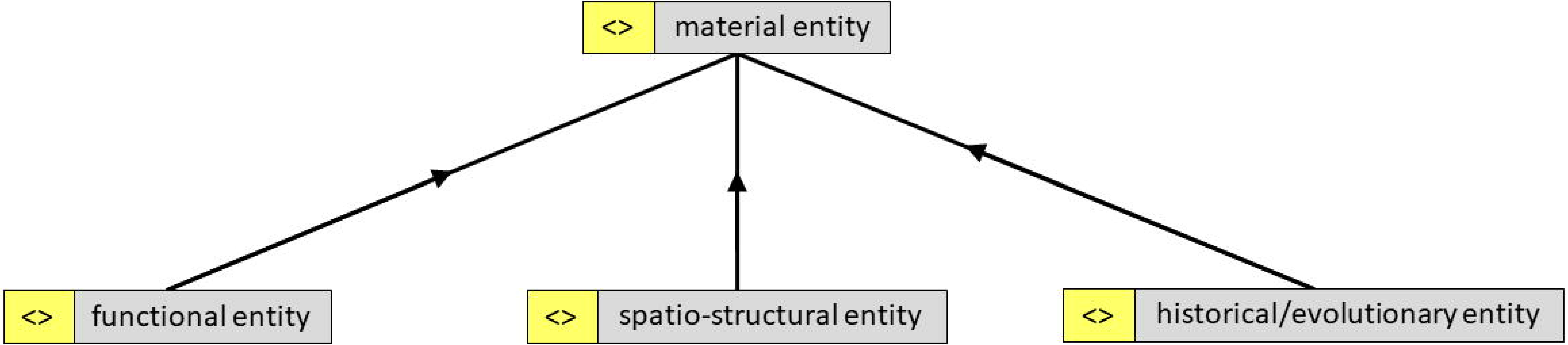
Fig 2. Different frames of reference within material entities. Arrows indicate an is_a relation property.

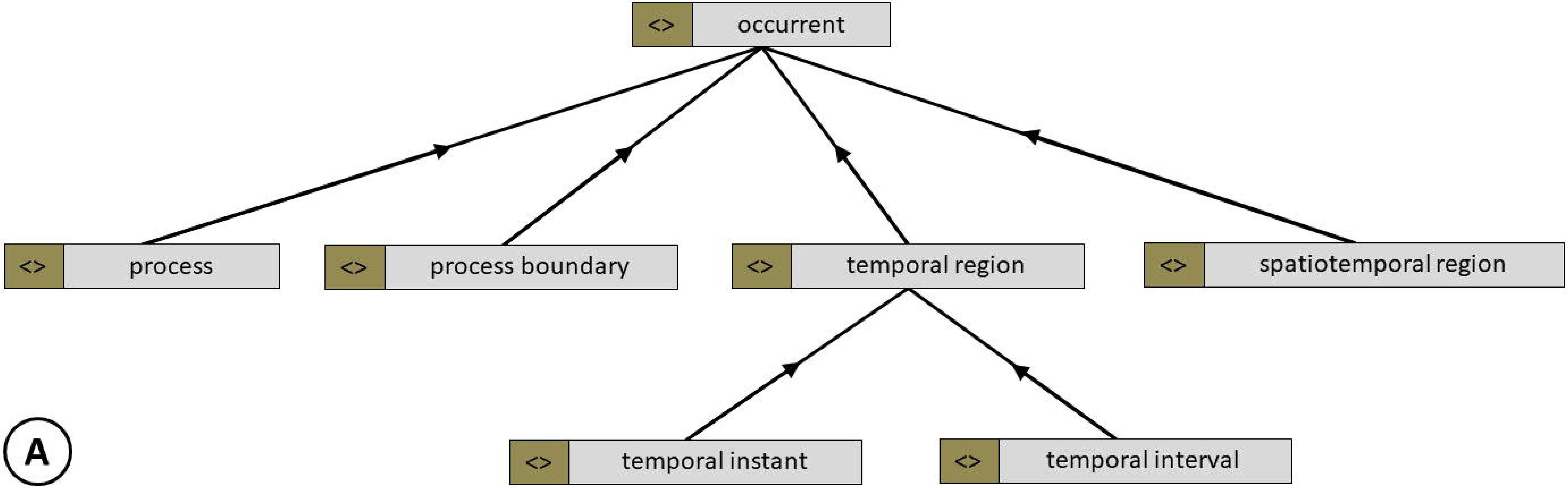
Fig 3. Distinct top-level categories of entities and their respective relations following BFO’s template. **A.** Top-level categories of occurrents, with their respective relations. Arrows indicate an *is_a* relation property. **B.** An exemplification of a relation between a given material entity, a given function, and a given process.

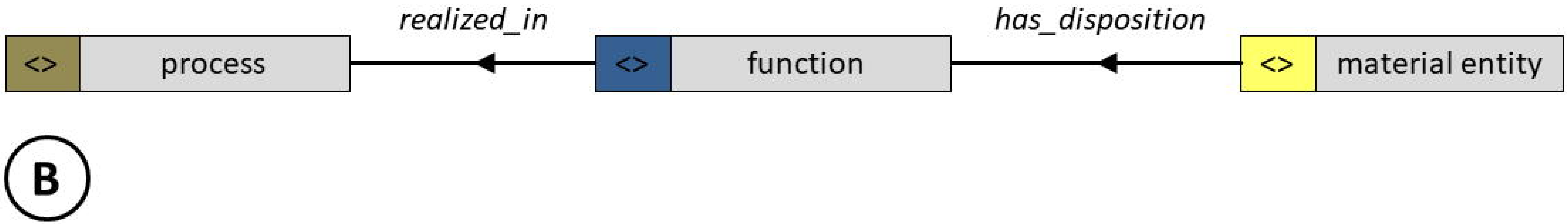
Fig 4. Simplified examples of entities and relations of ‘caste’ concepts modeled following distinct granular perspectives. **A**. An example of entities and relations of ‘caste’ concepts modeled following a CBB granular perspective. **B.** An example of entities and relations of ‘caste’ concepts modeled following a CFU granular perspective. **C.** An example of entities and relations of ‘caste’ concepts modeled following a CH/EU granular perspective. Entities marked with the “less-greater than” boxes of the same color belong to the same granular level.

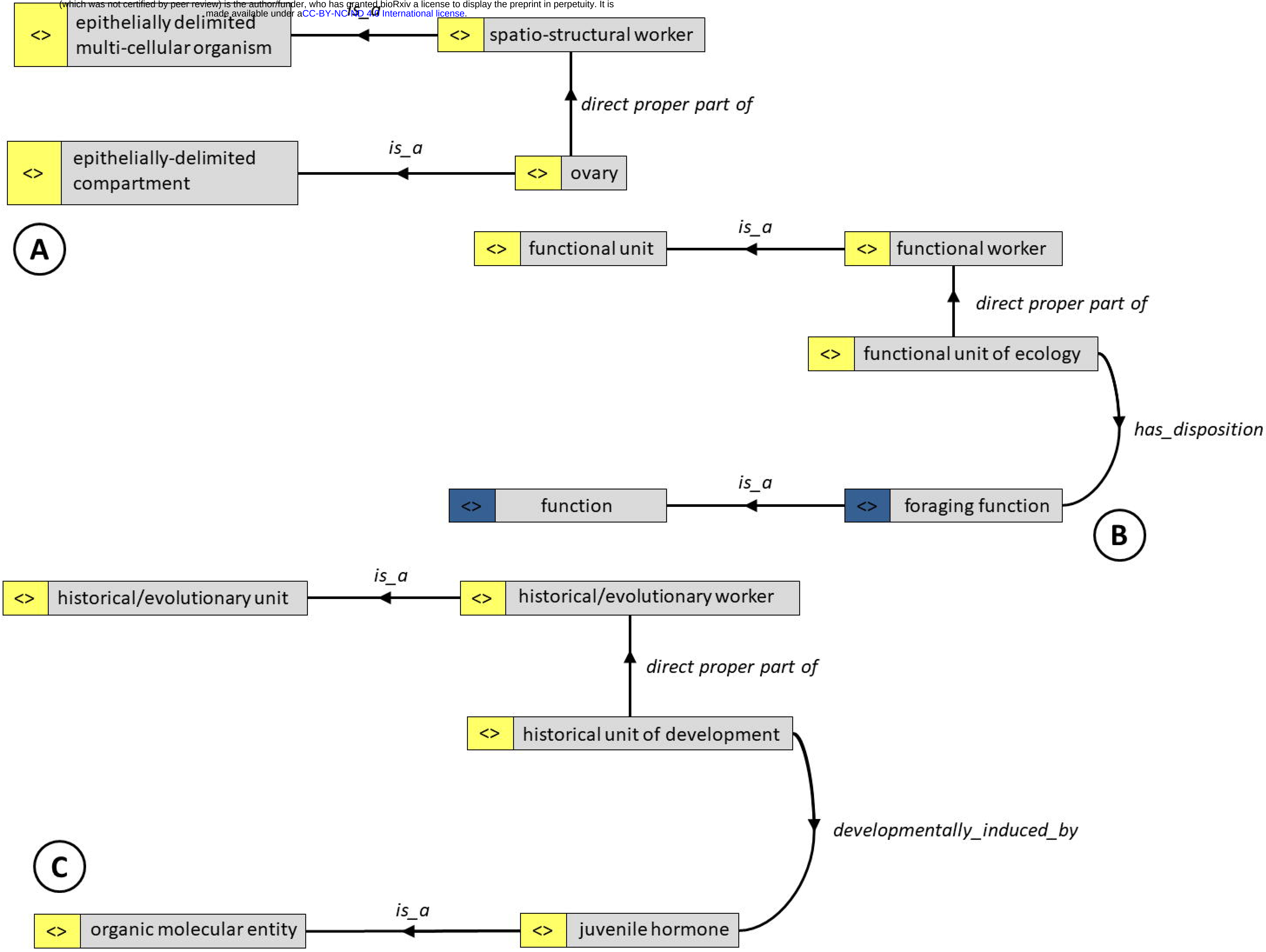
Fig 5. A simplified model of Vogt’s (2019) granularity framework, showing the integration among distinct granular perspectives through specific relations of granular representation. Some granular perspectives were omitted. CBB: Compositional Building Block; CBB-C: Compositional Building Block Cluster; F-BR: Function-based Representation; H/E-BR: Historical/Evolution-based Representation; FuncGranRep: has functional granular representation; Hist/EvGranRep: has historical/evolutionary granular representation. Adapted from Vogt (2019).


entity
occurrent
continuant
specifically dependent continuant
independent continuant
A
realizable entity
role
disposition
function
C
material entity
fiat object part
fiat object part aggregate
object aggregate
object with fiat object part aggregate
object
fiat object part cluster
fiat object part group
object group
object cluster
object with fiat object part cluster
object with fiat object part group
B

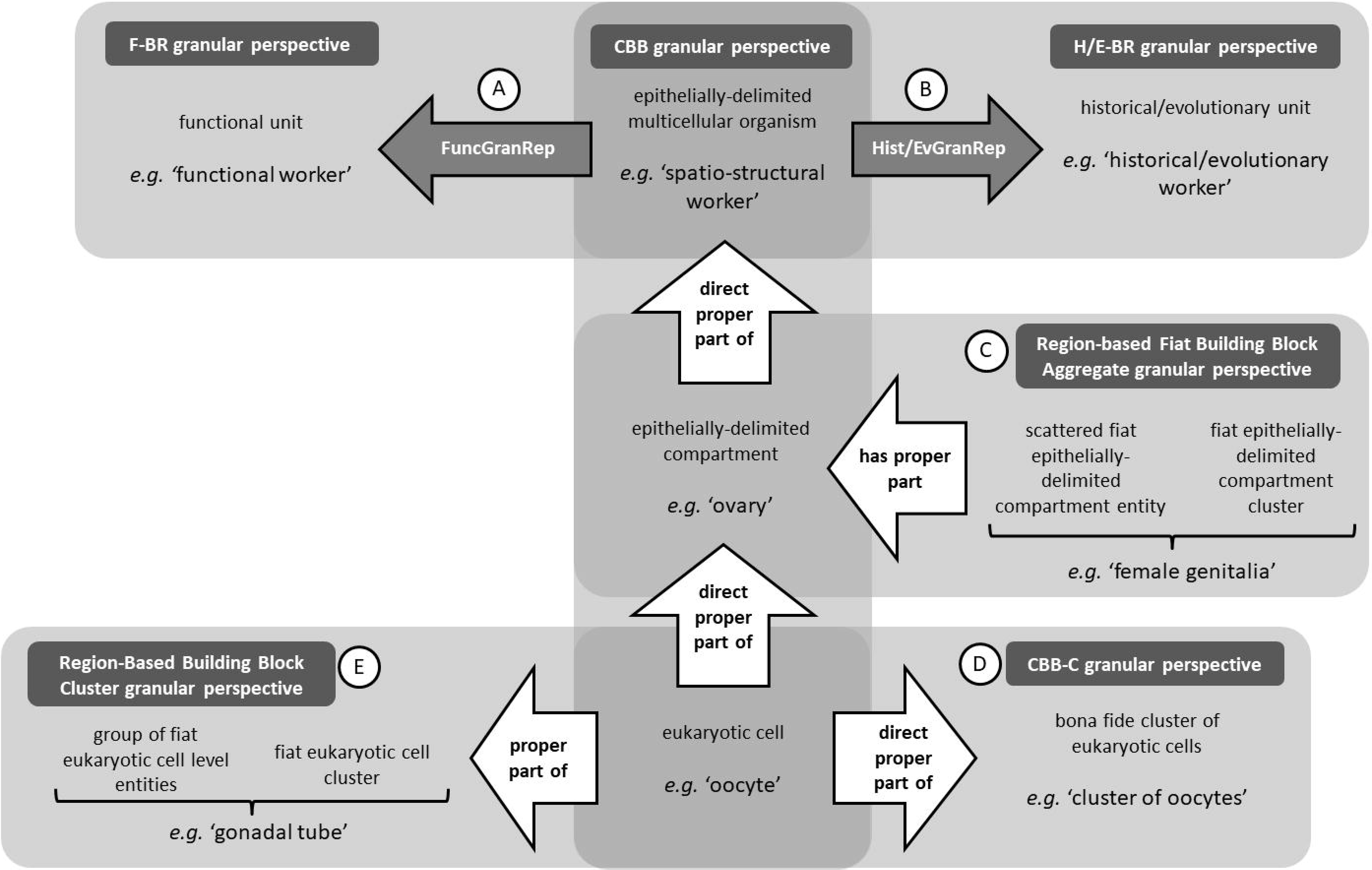

<> material entity
<> functional entity
<> spatio-structural entity
<> historical/evolutionary entity


occurrent
process
process boundary
temporal region
spatiotemporal region
temporal instant
temporal interval
A


process
realized_in
function
has_disposition
material entity
B



<>
epithelially delimited multi-cellular organism
is_a
spatio-structural worker
direct proper part of
epithelially-delimited compartment
is_a
ovary
A
functional unit
is_a
functional worker
direct proper part of
functional unit of ecology
has_disposition
function
is_a
foraging function
B
historical/evolutionary unit
is_a
historical/evolutionary worker
direct proper part of
historical unit of development
developmentally_induced_by
C
organic molecular entity
is_a
juvenile hormone


F-BR granular perspective
functional unit
e.g. 'functional worker'
A
FuncGranRep
CBB granular perspective
epithelially-delimited multicellular organism
e.g. 'spatio-structural worker'
B
Hist/EvGranRep
H/E-BR granular perspective
historical/evolutionary unit
e.g. 'historical/evolutionary worker'
direct proper part of
C
Region-based Fiat Building Block Aggregate granular perspective
epithelially-delimited compartment
e.g. 'ovary'
has proper part
scattered fiat epithelially-delimited compartment entity
fiat epithelially-delimited compartment cluster
e.g. 'female genitalia'
direct proper part of
Region-Based Building Block Cluster granular perspective
E
group of fiat eukaryotic cell level entities
fiat eukaryotic cell cluster
e.g. 'gonadal tube'
proper part of
eukaryotic cell
e.g. 'oocyte'
direct proper part of
D
CBB-C granular perspective
bona fide cluster of eukaryotic cells
e.g. 'cluster of oocytes'