INPRODAT Innovationszentrum für Prozessdatentechnik e.V.

Technical report no. 2020–B
original version

Multiscale modelling and simulation of physical systems as semiosis

---

Date: 21st June 2020

Authors: Horsch, M. T.; Chiacchiera, S.; Seaton, M. A.; Todorov, I. T.

Funding information:

- No external funds were accessed.

Accessibility:

- doi: 10.5281/zenodo.3902901
- https://zenodo.org/communities/inprodat/

Multiscale modelling and simulation of physical systems as semiosis

Martin Thomas HORSCH [a] Silvia CHIACCHIERA [a] Michael A. SEATON [a]
Ilian T. TODOROV [a]

[a] *UK Research and Innovation, STFC Daresbury Laboratory, Keckwick Ln, Daresbury, Cheshire WA4 4AD, United Kingdom*

Abstract. It is explored how physicalist mereotopology and Peircean semiotics can be applied to represent models, simulations, and workflows in multiscale modelling and simulation of physical systems within a top-level ontology. It is argued that to conceptualize modelling and simulation in such a framework, two major types of semiosis need to be formalized and combined with each other: Interpretation, where a sign and a represented object yield an interpretant (another representamen for the same object), and metonymization, where the represented object and a sign are in a three-way relationship with another object to which the signification is transferred. It is outlined how the main elements of the pre-existing simulation workflow descriptions MODA and OSMO, *i.e.*, use cases, models, solvers, and processors, can be aligned with a top-level ontology that implements this ontological paradigm, which is here referred to as mereosemiotic physicalism.

Keywords. Top-level ontology, semantic interoperability, simulation workflow, data provenance, semiotics.

1. Introduction

A model of the physical behaviour of a system is a sign: It represents the modelled system. Semiotics is the science of representation; a simulation that evaluates a model numerically is a semiotic process – a semiosis. By this semiosis, the meaning of the model as a sign realizes itself. In the same sense, modelling, understood as the process by which a model is designed and parameterized, is a semiosis. Each time that a numerical simulation is used to make predictions about a certain system or substance, various steps of cognition need to be undertaken: Typically, motivated by the purpose at hand, the real problem is simplified and tackled with an analytical model that is then converted into a numerical implementation, which produces an outcome addressing the initial question. In common practice, this process can consist of multiple iterations, repeated until a satisfactory understanding is attained. Each step can be a complex workflow itself, combining various models and solvers, ultimately all referring to the eventual use case.

Since the appearance of ontologies in computer science [1], there have been endeavours to develop ontologies for modelling and simulation to support the exchange of information both between people (communication) and between software (interoperability) [2,3]. This work aims at contributing to the description of simulation workflows by ontologies. It is concerned with the fundamentally ontological objective of characterizing what models and simulations *are* and how to describe this formally in a coherent way [3]. Solutions that have been proposed include ontologies targeted specifically at data technology for simulation-based engineering such as PhysSys [2] and, more recently, the Physics-based Simulation Ontology (PSO) [4], the Ontology for Simulation, Modelling, and Optimization (OSMO) [5], the Simulation Intent ontology [6], and the European Materials and Modelling Ontology (EMMO), the latter of which is still under development [7]; PSO is a domain ontology aligned with the Basic Formal Ontology (BFO) as a top-level ontology [8], whereas OSMO is a domain ontology aligned with the EMMO within a framework established by the European Materials Modelling Council (EMMC) on the basis of preceding standardization efforts [9,10].

This work is structured as follows: On the basis of mereosemiotic physicalism, an ontological paradigm characterized in Section 2, multiple possible solutions for addressing a series of challenging aspects of designing a top-level ontology for simulation-based engineering are outlined in Section 3. In Section 4, it is discussed how ontologies that are developed accordingly can support data technology in materials modelling; this is assessed on the basis of requirements from modelling and simulation practice as specified by standards and recommendations of the EMMC and associated projects. A conclusion is given in Section 5.

2. Mereosemiotic physicalism

The term *mereosemiotic physicalism* is employed here for what we consider to be the core of the ontological paradigm under investigation, following Ghedini *et al.* [7] while maintaining some freedom for variation, *cf.* Section 3. The following core tenets are stated to be constitutive of the paradigm:

- Physicalist materialism, *i.e.*, only that exists which can be conceived of as physically real in the actual world or a hypothetical world that does not differ fundamentally from ours and to which the same laws of physics apply.
- Mereotopology, *i.e.*, all that exists is spacetime, in the sense of conceivably being part of a universe with three spatial dimensions and the dimension of time [11,12]. One of the fundamental relations is proper parthood P such that $a \mathsf{P} b$ means "a is a proper part of b;" mereology becomes mereotopology through its extension by a predicate for connectedness [11].
- Semiotics following Peirce, by which signs (*i.e.*, representamina) engage in a dialectical relationship with represented objects through a process the elementary steps of which are conceptualized as triads [13,14,15,16]; accordingly, signification S, where $s \mathsf{S} o$ would mean "s is a sign for the object o," cannot stand in isolation, but requires a third (other) element to function. However, it is left open whether a dyadic signification relation can be inferred from the semiotic triad, *cf.* Section 3.2.
- Nominalism, by which only individuals (*i.e.*, concreta) exist. Quantification cannot be applied to classes (*i.e.*, universals, concepts, or abstracta as such),

only to individuals (*i.e.*, to instances or elements of a class), and only individuals, not classes, can be elements of a class.

This paradigm combines mereotopology and physicalism, *i.e.*, an approach that is immediately suitable for discussing materials and their properties, with semiotics grounded in a complex sign-object interaction that can capture intricate scenarios from modelling and simulation as well as their relation to experimental data.

In the EMMO [7,17], regions of spacetime that are connected topologically are referred to as items (*i.e.*, they are individuals of the class Item);[1] otherwise, they are Collection individuals, consisting of multiple connected components (Item individuals) which are related to the collection by membership, a subproperty of proper parthood, isMemberOf $\sqsubseteq$ P. In the formalization that mereotopology is given in the present work, connected regions of spacetime (*i.e.*, items) are permitted to have proper parts that are not connected (*i.e.*, collections). Individuals of both types, items and collections, can participate in semiosis in the role of the object. Physicalist materialism is well-aligned with a view of operating with signs (and of thought in general) as constituted by processes that are simultaneously logical, social, and material; much of Peirce's work can be read in this way. This approach contributes to semantic interoperability for infrastructures that deal with multiscale modelling and simulation of physical systems. Computer simulations are processes by which a sign (a model) is evaluated, yielding another sign (a different model or a computed property, *i.e.*, an interpretant), and data management needs to take the material preconditions of this process into account, since they constitute relevant metadata: Where was the simulation done, how were the input and the output stored and exchanged, *etc.*? Any platform architecture that aims at making models and simulation results findable, accessible, interoperable, and reusable (FAIR) must include such metadata to a certain extent [18,19]. Similarly, metadata on the provenance of sensory data are crucial for integrating "field devices 4.0" into model-driven process control [20]. In this view, such an approach seems to be well suited for its purpose.

An internal tension within the paradigm established by Ghedini *et al.* [7] stems from the fact that it is based on nominalism in combination with Peircean semiotics, while Peirce himself held a realist position, stating that "modern nominalists are mostly superficial men" and that "the conception of a pure abstraction is indispensable" [14]. However, as argued above, physicalist nominalism has aspects that are beneficial for the present purpose, and similarly, placing semiotics at the core of a conceptualization of dealing with models seems to be adequate; solving this tension requires flexibility and innovation, some of which can be found in the work on the EMMO [7,17]. In their approach, nominalism is subordinate to mereotopology and physicalism; *i.e.*, universals do not exist *because they are not spacetime*, and only spacetime exists. This is not "ostrich nominalism" [21], since it permits workarounds for including particulars that are similar to universals in many respects; *e.g.*, the *class* of all balloons filled with helium is not an individual, and hence, it does not exist. However, there is a collection that contains all

[1] The EMMO is a top-level ontology that, as this work is being redacted, has not yet been released in a stable version [17]; any remarks here refer to the version available in the git repository [7] by 28th February 2020.

these balloons as members – this collection is defined by the union of the regions of spacetime occupied by all the balloons. Moreover, the text "balloon filled with helium" exists and can act as a sign, so that it refers to the respective objects by semiosis. Both the collection and the sign can be identified with regions of spacetime. Hence, in the present approach, they can exist (as individuals), and their function parallels that of a class or a universal in certain ways.

The present conceptualization of semiotics includes two main categories of processes, which are included in the present ontology variants as subclasses of Semiosis: The first one is Interpretation, in an instance of which a sign s refers to an object o through its transposition into the other sign s'

$$i: \; s — o — s' \; [g_i], \tag{1}$$

which is called the interpretant, *i.e.*, a third element that also serves as a sign for the same object; *e.g.*, upon reading the phrase "balloon filled with helium," the reader may mentally visualize such an object. The phrase is the sign s, which is in this case *symbolic*, the internal visualization is the interpretant s' (in this case, *imaginary*), and the referenced balloon object o is *real* or may, alternatively, exist within a hypothetical scenario. In Relation (1), i is the interpretation instance, and g_i is its ground, in the Peircean sense [13,14]. The ground is an abstractum in Peirce, permitting the unification of multiple semioses, such as thoughts of different people, on the basis of a shared idea [13,14], *i.e.*, colloquially, a common ground; the inclusion of unification in the present nominalist approach is discussed in Section 3.4. This example is a special case where Peircean sign, interpretant, and object coincide with the Lacanian triad of the *symbolic*, the *imaginary*, and the *real*; typically, modelling and simulation follows a mode of reasoning where the sign and the interpretant are cognized as symbolic. However, there may be a more rigorous alignment between the symbolic-imaginary-real triad and the Peircean distinction between *symbols* (called *conventionals* in the EMMO [7]), *icons*, and *indices*. These are categories of signs which are, respectively, accessible to symbolic, imaginary, and real modes of cognition; *cf.* Balat [22] concerning the alignment of Peircean and Lacanian triads.

Following Peirce, the signification $s \, \mathsf{S} \, o$ and $s' \, \mathsf{S} \, o$ is meaningful only through its embedding in a triad and its context, including possible further triads, *e.g.*,

$$i, i', i'': \; s — o — s', \; s' — o — s'', \; s'' — o — s''' \; [g_i, g_i', g_i'']. \tag{2}$$

A sequence of interpretation processes is also itself an interpretation; *e.g.*, Relation (2) can be contracted to

$$i \circ i' \circ i'': \; s — o — s''' \; [g_i \circ g_i' \circ g_i''], \tag{3}$$

where the semioses are joined together by concatenation; *n.b.*, a semiosis process that only consists of interpretation steps can never depart from its initial object. Teleological reasoning is here represented by *telesis*, an interpretation

$$t: \; \tau — \alpha — \tau' \; [g_t], \tag{4}$$

where the sign is the telos τ, the object is the purposeful action α (the process by which the telos is pursued), and the interpretant τ' evaluates the outcome of the attempt; *e.g.*, this might be a revised objective that motivates the next step of action, or it might be a report on how successfully the aim was attained.

The second category of semiosis processes considered here, beside Interpretation, is Metonymization, which contributes a complementary aspect to the dialectics of the referent and the representamen. A metonymization is given by

$$m: \; o — s — o' \; [g_m], \tag{5}$$

where the object o and the sign s are related to the other object o', to which the same sign then also refers [23,24]; again, following Peirce, the signification $s \, \mathsf{S} \, o$ and $s \, \mathsf{S} \, o'$ can only function through a semiosis process that includes interpretation. As above, a sequence of metonymizations retains the quality of a metonymization. The Semiosis instance (*e.g.*, above, i or m) and the ground will be omitted from the notation where their existence is understood.

For details on the suggested implementation of this paradigm, the reader is referred to the ontology TTL file made available as supplementary material [25].

3. Variations within the paradigm

Mereosemiotic physicalism, defined above on the basis of ongoing work on the EMMO [7,17], permits describing modelling and simulation as semiosis. In this way, central concepts from simulation-based engineering data technology can be given a fundamental function in terms of entities defined in a top-level ontology; further domain ontologies can then be aligned with the top-level ontology [26], yielding a data infrastructure where models, data, and services can be integrated in accordance with the principles of FAIR data stewardship [18,27]. However, the core tenets of this ontological paradigm also raise issues that are non-trivial to address. They permit a variety of perspectives, some of which we aim to explore and comment below. In particular, this concerns 1) the relation between semiotics and physicalism; 2) the relation between triads and dyads in Peircean semiotics; 3) the role of objects from fictional or counterfactual scenarios and their participation in semiosis; 4) the equivalence of multiple instances or copies of the same symbol, data item, or simulation workflow.

3.1. Relation between semiotics and physicalism

Two variants will be considered:

- *Fundamentality of semiosis* (element $\mathcal{S}_1$), *i.e.*, the ontology is developed without assuming that the relation of represented and representing entities during semiosis can be expressed rigorously in physical terms. This includes conceptualizations where signification/semiosis and spatiotemporal parthood are both fundamental on an equal footing, or where symbolic reasoning is considered on the basis of formal logic only, without specifying any precise connection to the mereotopological characterization of spacetime.

- *Fundamentality of spatiotemporal parthood* (element $\mathcal{P}_1$): In this view, semiosis is a process, *i.e.*, a spatiotemporal region, in which any participating representamina are actually physically present; this participation is spatiotemporal parthood [7]. In a *primary* semiosis, defined here through the absence of a preceding semiosis from which the sign-object relation is carried over, the represented object needs to be physically present as well. In cases of *secondary* semiosis, the presence of the object is not required.

  Peirce states that every semiosis has a logical precursor; *i.e.*, in the present terms, that there is only secondary semiosis. Obversely, the EMMO prerelease version requires the spatiotemporal parthood of the object in every semiosis, implying that there is only primary semiosis [7]. Future versions of the EMMO might relax this very restrictive conceptualization.

Peirce's semiotics contains aspects of both elements: It is fundamentally nonphysicalist insofar as it admits the existence of universals; however, it requires a (possibly indirect) causal connectivity between the sign and the object. It distinguishes between semioses where the object is physically present and cases where the physical causal relation between the sign and the object may be indirect. The latter include references to hypothetical entities (permitted by Peirce) which by their nature cannot be a part of a non-hypothetical process, *cf.* Section 3.3.

3.2. Relation between triads and dyads

Two types of realizations of the paradigm can be distinguished on the basis of treating the semiotic triad either as fundamentally irreducible or as nonfundamental and reducible to a dyadic signification relation:

- *Irreducibility of the semiotic triad* (element $\mathcal{I}_2$). Following Peirce, mere dyadic signification of the type $s \ \mathsf{S} \ o$ does not occur; representation can only be realized in combination with a third element. Hence, the only pairwise relations that can be asserted in semiosis do not connect the sign directly to the object but, instead, the sign to the semiosis process, the object to the semiosis process, and the third element to semiosis process, respectively.
- *Reducibility by existential quantification* (element $\mathcal{E}_2$). Accordingly, where there is a triad $s — o — s'$, dyadic signification $s \ \mathsf{S} \ o$ and $s' \ \mathsf{S} \ o$ can be inferred. This is the approach implemented within the EMMO, which includes a relation hasSign between the object of reference and a sign [7,17].

3.3. Relation between fact and fiction

This relates to modelling and simulation of scenarios that are known not to have occurred, or of multiple scenarios that contradict each other, *e.g.*, by parameter variation, so that they cannot all be realized in the same universe or possible world. Ontologically, this reduces to the question whether the object of a factual semiosis (*e.g.*, of a simulation that is actually carried out) can be counterfactual. A rough classification is possible, differentiating approaches that permit a combination of fact and fiction from approaches that negate this possibility:

- *Inclusion of multiple modes of existence or multiple possible worlds* (element $\mathcal{M}_3$). Such approaches may be referred to as modal or Meinongian [28], since they include multiple modes in which objects can exist within the same knowledge base or ontology. This includes Peirce's semiotics which permits referring to hypothetical objects through signs that are not merely hypothetically, but factually present.

- *Negation of fictitious entities, or of the possibility to combine fact and fiction* (element $\mathcal{N}_3$). It is assumed that the referent of a sign needs to exist in the same reality as the sign itself; hence, either both are factual, or both are fictional. In particular, this includes the understanding according to which there is no secondary semiosis, so that the sign, the object, and the third element are all necessarily physically present (formalized as parthood) during semiosis [7]. Since something that exists factually cannot have a spatiotemporal part that only exists in fiction, it is impossible for a fictional object to be represented within an actual simulation workflow.

3.4. Relation between multiple copies

Dealing with multiple copies of the same data or metadata item is one of the most basic functions of data technology. It is a prerequisite for any exploitation of semantic interoperability in practice. But under what conditions are multiple signs (*e.g.*, models or simulation results), or multiple semioses (*e.g.*, simulations) to be regarded as the same individual, as equivalent or similar enough, or as manifestations of the same information content? How can a unification or subsumption of multiple entities under some shared identity be expressed when they occur in different parts of spacetime, *e.g.*, on different computers?

Peirce defines different types of signs for this purpose: "the word 'the' will usually occur from fifteen to twenty-five times on a page. It is in all these occurrences one and the same word, the same legisign. Each single instance of it is a Replica. The Replica is a Sinsign" [13]. There, the legisign is either a universal, or if it is understood as a particular, it does not have a clearly specified location as a part of spacetime. Possible approaches for implementing this within the present paradigm may include:

- *Unification by universals* (element $\mathcal{U}_4$). Multiple concreta can share a feature by partaking in the same abstractum. This is closest to Peirce's realism. For an implementation that is compatible with the nominalist paradigm, a variety of solutions can be thought of; *e.g.*, classes, sets, or collections may be defined that contain all individuals that are similar in a certain respect, or an equivalence relation may be employed to state that its subject and object are replicas of the same item. What these approaches have in common is that it is externally posited whether two items are equivalent or not: The strings "the" and "the" are the same word because the knowledge base (or, following Ghedini *et al.* [7], "the ontologist") states it. It is not possible for one interpreter to recognize them as the same and for another to believe them to be different.

- *Absence of unification* (element $\mathcal{A}_4$). If whatever exists is spacetime, and exists as spacetime, different regions of spacetime by definition cannot be the same; accordingly, "the" and "the" are just different physical objects, one of which is printed more to the left, while the other is printed more to the right. This is most in line with an ontology that prioritizes mereotopology over semiotics, including the pre-release version of the EMMO [7].
- *Unification by semiosis* (element $\mathcal{S}_4$). This solution proposes that assessing the equivalence, correctness, or validity of signs, processes, *etc.*, is a process of pattern matching, which is a sort of interpretation. Thereby, the sign is a pattern, the object is the item that is matched against the pattern, and the interpretant is the outcome of the pattern matching process, which may be acceptance, rejection, or any other assessment of how well the object matches the pattern. If two objects match the same pattern, they are equivalent in a certain respect; however, this is not externally posited, but subject to a process of semiosis, and multiple interpreters may disagree.

3.5. Design space for possible top-level ontologies

Based on the core tenets of the paradigm from Section 2 and a combination of elements from Sections 3.1 to 3.4, appropriately chosen, a variety of strategies can be followed for the design of a top-level ontology for modelling and simulation in engineering data technology. The EMMO pre-release version [7] is best described as a combination of the elements $\mathcal{P}_1\mathcal{E}_2\mathcal{N}_3\mathcal{A}_4$. Peirce [13,14] permits a reading that positions him comparably closely to the combination $\mathcal{S}_1\mathcal{I}_2\mathcal{M}_3\mathcal{U}_4$, as far as such a claim may be upheld for any ontology that follows nominalism rather than realism. The choices on how to deal with each of the four discussed challenges can be made independently; there are no combinations of elements that appear to be impossible to reconcile with each other. Therefore, the entire product space

$$\{\mathcal{P}_1, \mathcal{S}_1\} \times \{\mathcal{E}_2, \mathcal{I}_2\} \times \{\mathcal{M}_3, \mathcal{N}_3\} \times \{\mathcal{A}_4, \mathcal{S}_4, \mathcal{U}_4\} \tag{6}$$

is accessible, providing a landscape of possible types of top-level ontologies within the paradigm of mereosemiotic physicalism.

Within this design space, the combination $\mathcal{P}_1\mathcal{I}_2\mathcal{M}_3\mathcal{S}_4$ is particularly advantageous for applications in platform and infrastructure development for multiscale modelling and simulation of physical systems: 1) Fundamentality of spatiotemporal parthood (element $\mathcal{P}_1$) situates models and materials straightforwardly within the same framework and encourages the user to consider where and how exactly data are stored and simulations are carried out. 2) The irreducibility of the semiotic triad (element $\mathcal{I}_2$), where each semiosis is a part of a chain of logically connected semioses and becomes meaningful only through this context, has its direct correspondence in the necessity to provide contextual and provenance information as metadata to make models and simulation data FAIR [18,27]. Where data are obtained by simulation, this requires a formal description of simulation workflows [5,10]; cf. Section 4 on the expressibility of workflow descriptions by chains of semioses in terms of triads. 3) By admitting the existence of entities from multiple possible worlds within the same knowledge base (element $\mathcal{M}_3$), it can be stated that a simulation is *factually done* on the basis of a *factually existing* model of a *fictional scenario* that has not happened and hopefully will not happen, *e.g.*, a severe accident. Similarly, a process can be optimized by parameter variation on the basis of a model, without implying that all the combinations that have been modelled will actually be built in the real world. 4) Unification by semiosis (element $\mathcal{S}_4$) encourages the user to state clearly in what way two instances were determined to be replicas or manifestations of the same thing, *e.g.*, by describing how they match the same pattern or share the same common ground. It is thus taken into account that it is often ambiguous whether two implementations of "the same" model or simulation workflow are really the same. The round-robin study of simulation scenarios by Schappals *et al.* [29] demonstrates this, finding significant deviations between results proceding from what might be described as equivalent workflows, even for relatively benign cases.

As a proof of concept, a demonstrator implementation of $\mathcal{P}_1\mathcal{I}_2\mathcal{M}_3\mathcal{S}_4$ that remains as close as possible to the EMMO pre-release version, named EMMO-PIMS, is provided here as supplementary material [25]; *cf.* Figs. 1 and 2 for a selected part of the relation and class hierarchies from EMMO-PIMS.

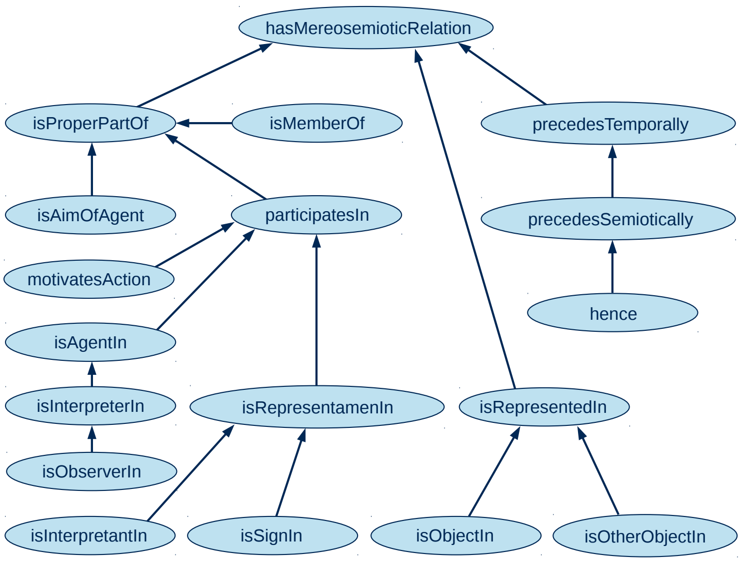

Figure 1. Relation hierarchy (partial) for the present demonstrator top-level ontology implementation of the element combination $\mathcal{P}_1\mathcal{I}_2\mathcal{M}_3\mathcal{S}_4$. The arrows represent the transitive reduction of rdfs:subPropertyOf, *i.e.*, an arrow $r \rightarrow r'$ implies $r \sqsubseteq r'$; see the supplementary TTL file [25] for a description of the semantics of these relations as well as the associated domains and ranges.

4. Application to simulation workflows

Many typical features of modelling and simulation workflows, as well as associated data and metadata, can be understood as instances of semiosis, or as being involved in semiosis through Peircean triads. The paradigm of mereosemi-

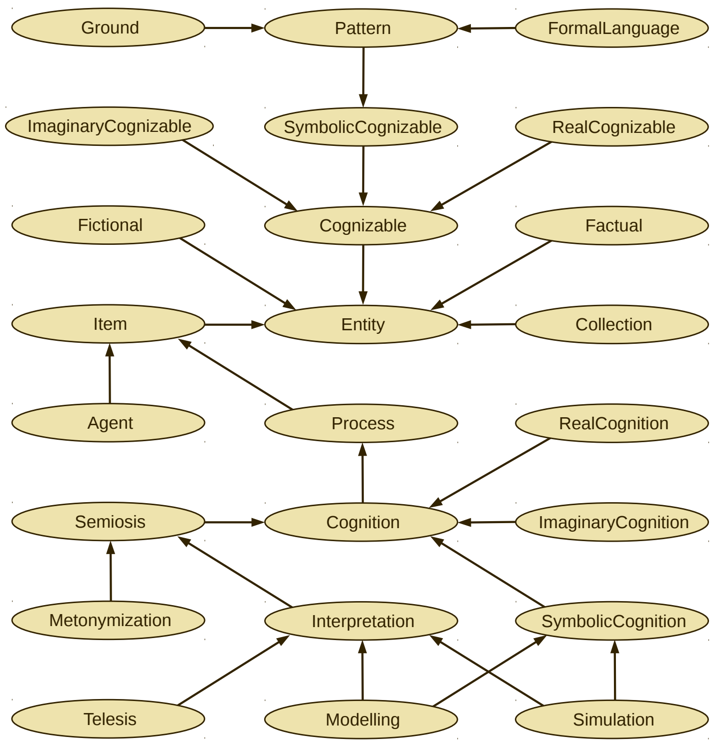

Figure 2. Class hierarchy (partial) for the present demonstrator top-level ontology implementation of the element combination $\mathcal{P}_1\mathcal{I}_2\mathcal{M}_3\mathcal{S}_4$. The arrows represent the transitive reduction of rdfs:subClassOf, *i.e.*, an arrow $c \rightarrow c'$ implies $c \sqsubseteq c'$; see the supplementary TTL file [25] for a description of the semantics of these classes as well as the associated rules and relations.

otic physicalism can therefore be applied to workflows in multiscale modelling and simulation of physical systems. According to the semantic assets concerning simulation workflows that have emerged from the ongoing standardization efforts of the EMMC, *i.e.*, the Review of Materials Modelling (RoMM) [9], the semi-formalized graph notation MODA [10], and the domain ontology OSMO [5], a workflow description is an arrangement of instances of four types of sections (*i.e.*, description elements): Use cases, models, solvers, and processors [5,10].

4.1. The use case

A use case is a product, process, or system of interest, posing a challenge that constitutes the *telos* for a modelling and simulation workflow. The definition of a use case following MODA [10] and OSMO [5] is relatively close to practice; it can be compared to the concept of "simulation intent" from Nolan *et al.* [30]. At a more aggregated level, the use case integrates with the concepts of the business case, the industrial case, and the translation case as defined in recommendations issued by the EMMC and associated projects, in particular the EMMC Translation Case Template [31] and the Materials Modelling Translation Ontology [32]. A formalization of the use case will typically consist of a step of metonymization

$$m_1 : \; o_1 — \tau_1 — o_2 \; [g_{m,1}], \tag{7}$$

and an associated telesis, *i.e.*, an interpretation

$$t : \; \tau_1 — i_1 — \tau_2 \; [g_t], \tag{8}$$

cf. Relation (4). There, the use case is an object o_1 to which the telos τ_1 refers, while it also refers metonymically to object o_2, the system that is to be modelled and simulated. In Relation (8), τ_1 is the sign, the object of telesis is the action motivated by it, *e.g.*, a simulation (which is itself an interpretation process, denoted here by i_1), and the interpretant is τ_2.

If the system that is modelled and simulated is identical with the use case, $o_1 = o_2$, the first step is not needed. In most cases, however, the system is a part or specific aspect of the use case, requiring metonymization as in Relation (7); it may even be a comparably very simple or small part, *e.g.*, if the use case is a production process or a value-added chain that is to be optimized.

4.2. The model

A model can occur as a sign in a *simulation* process, *i.e.*, an interpretation

$$i_1 : \; \mu_1 — o_2 — \mu_2 \; [g_{i,1}], \tag{9}$$

where μ_1 is the model, and the object is the simulated system o_2, which is related to a use case by metonymization, *cf.* Relation (7); the interpretant μ_2, also acting as a representamen for o_2, is given by the simulation outcome.

Nobody does a simulation by accident; since a simulation is a conscious purposeful action, it needs to be motivated by an aim τ_1, which also acts as a representamen both for the simulation process itself, *cf.* Relation (8), as well as for the the use case and the modelled system, *cf.* Relation (7). This relation also serves as the initial point for the *modelling* process by which the model is obtained, which may include model design, parameterization, selection from a database, *etc.*; accordingly, modelling (model development) is an interpretation

$$i_2 : \; \tau_1 — o_2 — \mu_1 \; [g_{i,2}], \tag{10}$$

where the model, as representamen obtained from the process, is the interpretant.

4.3. The solver

The solver s is a numerically processable version of a model, *e.g.*, of the model μ_1; to function as intended, they need to represent the same object,[2] so that the implementation and/or technical setup of the solver is given by the interpretation

[2]Note that both μ_1 and s here represent the object o_2. They might even be identical, $\mu_1 = s$, in which case Relation (11) is not needed. However, for an alignment with MODA and OSMO, it is preferable to maintain a distinction between the model μ_1 and the solver s that encapsulates its computational representation; if this distinction from MODA and OSMO is to be emphasized, the *simulation* triad from Eq. (9) can equally properly be denoted by $s — o_2 — \mu_2$.

$$i_3 : \ \mu_1 — o_2 — s \ [g_{i,3}]. \tag{11}$$

If the solver is used to conduct the simulation i_1, *cf.* Relation (9), it must be mereotopologically contained by i_1 through spatiotemporal proper parthood

$$s \ \mathsf{P} \ i_1. \tag{12}$$

4.4. The processor

The simulation result may subsequently be aggregated or transformed, *e.g.*, yielding a higher-level model of the same object by interpretation

$$i_4 : \ \mu_2 — o_2 — \mu_3 \ [g_{i,4}]. \tag{13}$$

This is typically taken care of by a processor, a simulation workflow element within MODA/OSMO that facilitates coupling and linking. Alternatively, for the result μ_2 of the first simulation step (where o_2 was simulated) to be applied to a related, but not identical object o_3 in a subsequent simulation step, a metonymization

$$m_2 : \ o_2 — \mu_2 — o_3 \ [g_{m,2}] \tag{14}$$

needs to occur first, which might also be supported by a processor.[3] In both cases, the processor π relates to the semiosis in the same way as the solver relates to the simulation, *i.e.*, by spatiotemporal proper parthood, $\pi \ \mathsf{P} \ i_4$ or $\pi \ \mathsf{P} \ m_2$, respectively. Such steps are abundant wherever complex simulation workflows are employed in practice [5]; *e.g.*, in COSMO-RS [33] or COSMO-SAC [34], a σ-profile (model) is obtained from a simulation of a single molecule and can then be applied to fluid phase equilibria [35]. This can only be conceptualized by combining interpretation (simulation, with a solver) and metonymization (with a processor).

5. Conclusion

From the present discussion, which is preliminary in nature, it can be concluded that physicalism, mereotopology, Peircean semiotics, and nominalism constitute a coherent ontological paradigm, mereosemiotic physicalism, that is highly suitable for applications in simulation-based engineering. The pre-release version of the EMMO has certain limitations, outlined above, that (in its present formulation) restrict its applicability to typical semioses in modelling and simulation practice; addressing this may require a dedicated effort. Building on this, is needs to be further investigated how exactly MODA and OSMO should best be aligned with the EMMO in detail, and to what extent a shared top-level ontology that implements mereosemiotic physicalism can contribute to achieving platform and service interoperability in practice, involving virtual marketplaces [26], model databases [36,37], simulation platforms [38], and other infrastructures.

[3] In MODA and OSMO, solvers and processors are distinguished by their roles; it is not necessary for these elements to correspond to separate software packages [5,10].

Simulation workflow descriptions from engineering practice may be very complex and consist of a large number of sections, *cf.* the MODA workflow graph examples from RoMM [9] and the scenario documented in previous work on OSMO [5]. Nonetheless, characterizing the MODA/OSMO sections in terms of a top-level ontological formalism, *cf.* Section 4, may contribute to aligning such descriptions rigorously with the EMMO, supporting platform interoperability and provenance documentation for data originating from simulation workflows.

Acknowledgment. This work was facilitated by activities of the Innovation Centre for Process Data Technology (Inprodat e.V.). No external funds were accessed.

Supplementary material. The *EMMO-based Physicalist Interpretation of Modelling and Simulation* (EMMO-PIMS) is a demonstrator implementation of a top-level ontology based on mereosemiotic physicalism, variant $\mathcal{P}_1\mathcal{I}_2\mathcal{M}_3\mathcal{S}_4$, *cf.* Section 3.5. EMMO-PIMS is available in TTL format, supplementing this work [25].

References

[1] T.R. Gruber, Toward principles for the design of ontologies used for knowledge sharing?, *Int. J. Hum.-Comp. St.* **43**(5–6) (1995), 907–928.

[2] P. Borst, H. Akkermans and J. Top, Engineering ontologies, *Int. J. Hum.-Comp. St.* **46**(2–3) (1997), 365–406.

[3] C. Turnitsa, J.J. Padilla and A. Tolk, Ontology for modeling and simulation, in: *Proceedings of the 2010 Winter Simulation Conference*, B. Johansson, S. Jain, J. Montoya Torres, J. Hugan and E. Yücesan, eds, IEEE, Piscataway, New Jersey, 2010, pp. 643–651.

[4] H. Cheong and A. Butscher, Physics-based simulation ontology: An ontology to support modelling and reuse of data for physics-based simulation, *J. Eng. Design* **30**(10–12) (2019), 655–687.

[5] M.T. Horsch, C. Niethammer, G. Boccardo, P. Carbone, S. Chiacchiera, M. Chiricotto, J.D. Elliott, V. Lobaskin, P. Neumann, P. Schiffels, M.A. Seaton, I.T. Todorov, J. Vrabec and W.L. Cavalcanti, Semantic interoperability and characterization of data provenance in computational molecular engineering, *J. Chem. Eng. Data* **65**(3) (2020), 1313–1329.

[6] F. Boussuge, C.M. Tierney, H. Vilmart, T.T. Robinson, C.G. Armstrong, D.C. Nolan, J.-C. Léon and F. Ulliana, Capturing simulation intent in an ontology: CAD and CAE integration application, *J. Eng. Design* **30** (2019), 688–725.

[7] E. Ghedini, Y. Bami, J. Friis, A. Hashibon, G.J. Schmitz, D. Toti and G. Goldbeck, European Materials and Modelling Ontology (EMMO), pre-release, 2020, https://github.com/emmo-repo/ and https://emmc.info/emmo-info/; date of access: 28th February 2020.

[8] R. Arp, B. Smith and A.D. Spear, *Building Ontologies with Basic Formal Ontology*, MIT Press, Cambridge, Massachusetts, 2015.

[9] A.F. De Baas (ed.), *What Makes a Material Function? Let me Compute the Ways*, EU Publications Office, Luxembourg, 2017.

[10] Materials modelling: Terminology, classification and metadata, CEN workshop agreement, 17284, CEN-CENELEC Management Centre, Brussels, 2018.

[11] A.C. Varzi, Parts, wholes, and part-whole relations: The prospects of mereotopology, *Data Knowl. Eng.* **20**(3) (1996), 259–286.

[12] B. Smith, Mereotopology: A theory of parts and boundaries, *Data Knowl. Eng.* **20**(3) (1996), 287–303.

[13] C.S. Peirce, *Logic as semiotic: The theory of signs*, in: *Philosophical Writings of Peirce*, J. Buchler, ed., Dover, New York, 1955, pp. 98–119.

[14] C.S. Peirce, *Peirce on Signs*, University of North Carolina Press, 1991.

[15] J. Zeman, The esthetic sign in Peirce's semiotic, *Semiotica* **19**(3–4) (1977), 241–258.

[16] T.L. Short, *Peirce's Theory of Signs*, Cambridge University Press, 2007.

[17] G. Goldbeck, E. Ghedini, A. Hashibon, G.J. Schmitz and J. Friis, A reference language and ontology for materials modelling and interoperability, in: *Proceedings of the NAFEMS World Congress 2019, Québec*, NAFEMS, Knutsford, UK, 2019, p. NWC_19_86.

[18] B. Mons, *Data Stewardship for Open Science*, CRC, Boca Raton, Florida, 2018.

[19] B. Schembera and D. Iglezakis, The genesis of EngMeta: A metadata model for research data in computational engineering, in: *Metadata and Semantic Research*, E. Garoufallou, F. Sartori, R. Siatri and M. Zervas, eds, Springer, Cham, 2019, pp. 127–132.

[20] M. Maiwald, Integrated and networked systems and processes: A perspective for digital transformation in thermal process engineering, *ChemEngineering* **4**(1) (2020), 15.

[21] D.H. Mellor and A. Oliver, *Properties*, Oxford University Press, 1997.

[22] M. Balat, Le musement, de Peirce à Lacan, *Rev. Int. Philos.* **46**(180) (1992), 101–125.

[23] C. Paradis, *Metonymization: A key mechanism in semantic change*, in: *Defining Metonymy in Cognitive Linguistics: Towards a Consensus View*, R. Benczes, A. Barcelona and F.J. Ruiz de Mendoza Ibáñez, eds, John Benjamins, Amsterdam, 2011, pp. 61–88.

[24] M. Brdar and R. Brdar-Szabó, *Translating (by means of) metonymy*, in: *Cognitive Linguistics and Translation: Advances in Some Theoretical Models and Applications*, A. Rojo and I. Ibarretxe Antuñano, eds, De Gruyter Mouton, Berlin, 2013, pp. 199–226.

[25] M.T. Horsch, S. Chiacchiera, M.A. Seaton and I.T. Todorov, EMMO-based Physicalist Interpretation of Modelling and Simulation (EMMO-PIMS), 2020, https://www.molmod.info/semantics/emmo-pims.ttl.

[26] M.T. Horsch, S. Chiacchiera, M.A. Seaton, I.T. Todorov, K. Šindelka, M. Lísal, B. Andreon, E. Bayro Kaiser, G. Mogni, G. Goldbeck, R. Kunze, G. Summer, A. Fiseni, H. Brüning, P. Schiffels and W.L. Cavalcanti, Ontologies for the Virtual Materials Marketplace, *Künstl. Intell.* (2020). doi:10.1007/s13218-020-00648-9.

[27] J. Bicarregui, Building and sustaining data infrastructures: Putting policy into practice, Technical Report, 2016. doi:10.6084/m9.figshare.4055538.v2.

[28] F. Berto and M. Plebani, *Ontology and Metaontology*, Bloomsbury, London, 2015.

[29] M. Schappals, A. Mecklenfeld, L. Kröger, V. Botan, A. Köster, S. Stephan, E.J. García, G. Rutkai, G. Raabe, P. Klein, K. Leonhard, C.W. Glass, J. Lenhard, J. Vrabec and H. Hasse, Round robin study: Molecular simulation of thermodynamic properties from models with internal degrees of freedom, *J. Chem. Theory Comput.* **13**(9) (2017), 4270–4280.

[30] D.C. Nolan, C.M. Tierney, C.G. Armstrong and T.T. Robinson, Defining simulation intent, *Comp.-Aid. Design* **59** (2015), 50–63.

[31] Europan Materials Modelling Council (EMMC), EMMC Translation Case Template, 2017, https://emmc.info/emmc-translation-case-template/; date of access: 21st March 2020.

[32] M.T. Horsch, S. Chiacchiera, M.A. Seaton, I.T. Todorov, B. Schembera, P. Klein and N. Konchakova, Pragmatic interoperability and translation of industrial engineering problems into modelling and simulation solutions, 2020, in preparation.

[33] A. Klamt, The COSMO and COSMO-RS solvation models, *WIREs Comput. Mol. Sci.* **8**(1) (2018), e1338.

[34] C.-M. Hsieh, S.-T. Lin and J. Vrabec, Considering the dispersive interactions in the COSMO-SAC model for more accurate predictions of fluid phase behavior, *Fluid Phase Equilib.* **367** (2014), 109–116.

[35] R. Fingerhut, W.-L. Chen, A. Schedemann, W. Cordes, J. Rarey, C.-M. Hsieh, J. Vrabec and S.-T. Lin, A comprehensive assessment of COSMO-SAC models for predictions of fluid phase equilibria, *Ind. Eng. Chem. Res.* **56**(35) (2017), 9868–9884.

[36] Å. Ervik, A. Mejía and E.A. Müller, Bottled SAFT: A web app providing SAFT-γ Mie force field parameters for thousands of molecular fluids, *J. Chem. Inform. Model.* **56**(9) (2016), 1609–1614.

[37] S. Stephan, M.T. Horsch, J. Vrabec and H. Hasse, MolMod – an open access database of force fields for molecular simulations of fluids, *Mol. Sim.* **45**(10) (2019), 806–814.

[38] A. Ribes and C. Caremoli, Salomé platform component model for numerical simulation, in: *31st Annual International Computer Software and Applications Conference*, Vol. 2, C.K. Chang, ed., IEEE, Los Alamitos, California, 2007, pp. 553–564.