Proceedings of the 2012 Winter Simulation Conference
C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher, eds

SEMIOTICS, ENTROPY, AND INTEROPERABILITY OF SIMULATION SYSTEMS – MATHEMATICAL FOUNDATIONS OF M&S STANDARDIZATION

Andreas Tolk

Department of Eng. Mgmt. & Sys. Eng.
Old Dominion University
Norfolk, VA 23529, USA

Saikou Y. Diallo
Jose J. Padilla

Virginia Modeling Analysis & Simulation Ctr.
Old Dominion University
Suffolk, VA 23435, USA

ABSTRACT

Semiotics identifies which symbols are used (syntax), what the meaning of these symbols is (semantics), and what the intention of using symbols is (pragmatics). These ideas have already been mapped to integratability of networks, interoperability of simulations, and composability of models for modeling and simulation applications. New research on model theory and algorithmic information theory support this viewpoint. Applying the finding of mathematics allows to define three different entropies: syntactical entropy that measures the variety of data representation, semantic entropy that measures the variety of data interpretation, and pragmatic entropy that measures the variety of data utilization. The paper shows the interconnection between these ideas and their implication for interoperability challenges: standards are needed on all levels to ensure meaningful interoperation, but their application reduces the interoperability space of federated solutions to the intersection of models, not to the union of models as often assumed in naïve approaches.

1 INTRODUCTION

What is modeling and simulation (M&S)? What are our philosophical and mathematical foundations we use to make the case for M&S? Do our recommended practices and standards make sense? These are fundamental questions that have not been addressed in sufficient detail and academic rigor. One could argue this is so because of the success story of globally distributed M&S federations in support of training and education, in particular in the military and defense application domain, did not give the M&S community time to "sit back and think." Another reason could be the success of the application of M&S techniques that we forgot to think about the theoretical implications of those techniques. There were too many practical problems to solve to reflect on the philosophical and scientific underpinnings of what was being done. We may have been like the proverbial lumberjack who had no time to sharpen his ax – or consider using a chain saw – because he had to cut down so many trees.

As a result, the philosophical underpinnings for M&S as a discipline have been discussed in the philosophy of science community (Frigg and Reiss 2009, Winsberg 2010), but are still in their infancy in the M&S community itself. A good starting point for bridging this gap is to look at existing works in related domains and examine whether and where they can be applied in M&S. Within this paper, we will address the previously posed questions using independently developed work on ontology, epistemology, semiotics, model theory, algorithmic information theory and several entropy measures can be systematically applied to M&S and thus be brought into the M&S body of knowledge. Further, we show that the applications on these concepts in M&S have important implications to composability and interoperability leading to a possible paradigm shift in applying M&S interoperability standards.

978-1-4673-4782-2/12/$31.00 ©2012 IEEE

2 RELATED WORK, DOMAINS, AND TERMS

In this section, related work on ontology, semiotics, and mathematics will be introduced in the context of their contribution to the topic of this paper. The use of ontologies to capture the concepts of application domains as well as the concepts we use to capture them has been the topic of several scholarly papers, such as the papers captured in (Tolk and Miller 2011). Semiotics identifies which symbols are used (syntax), what the meaning of these symbols is (semantics), and what the intention of using symbols is (pragmatics). These ideas have already been mapped to integratability of networks, interoperability of simulations, and composability of models for modeling and simulation applications. New research on model theory and algorithmic information theory support this viewpoint. This section introduces a summary of related work, domains, and terms. This list is not intended to be exclusive or complete but represents examples that can easily be extended.

2.1 Ontology and Epistemology

Turnitsa, Padilla, and Tolk (2010) introduced some ontological and epistemological interpretations for simulationists. The view expressed in their paper has been significantly influenced by the idea of epistemological and ontological preconceptions or worldviews from the perspective of problem solving, as presented by Bozkurt, Padilla, and Sousa-Poza (2007), as M&S is primarily perceived as a discipline supporting problem solving. One of the challenges that need to be overcome is that these terms are often overloaded. In the context of this paper, the terms ontology and epistemology are used as follows:

Ontology *is the study of being or what exists concerned with questions such as what are objects, their relations, and their properties.*

As such, the term is used more general than addressing information technology tools, such as Protégé, that focus on computer engineering views of ontology as a way to model within the ontological spectrum, as proposed by Silver (et al. 2007). However, ontology for M&S must answer the question of what propertied and associated concepts we use to express what exists in our models and simulations, which includes computational representations.

Epistemology *is the study of knowledge concerned with questions such as what are the necessary and sufficient conditions of knowledge and how true beliefs are justified.*

Ontology and epistemology are linked to one another in that an ontological position about something is linked to what we considered to be knowledge about that something. As such, epistemological and ontological foundations use and expand on one another. In this sense, Epistemology for M&S must answer questions related to the type of knowledge models and simulations generate. Ontology must answer questions about the frame of references under which knowledge is generated. Epistemology and ontology provide guidance to study semiotics in M&S by providing a platform of discussion on how to deal with syntax, semantics, and pragmatics issues in M&S, specially the symbols, relations, and functions we use to capture a phenomenon so it can be simulated using a computer and how we establish truth - at the model and simulation levels of abstraction - for a particular purpose.

2.2 Semiotics: Syntax, Semantics, and Pragmatics

Traditionally, semiotics addresses three interrelated fields, namely how a sign is defined by using a set of symbols (syntax), what the meaning of each symbol is (semantics), and what the intended use of the symbol is (pragmatics). Ögden and Richards (1923) introduced the semiotic triangle to evaluate the question on meaning and why we have problems to understand each other, even when we are describing the same referent in the real world. The answer was easy: although we may use the same real world referent, we build internal concepts to describe and understand it. We conceptualize based on our sensors and on a priori means and infer based on inductive and deductive processes. Then we use symbols to represent our symbols. While we assume that we represent the real world referent with the symbol, we actually represent the concept we have of this referent. However, on the communication level, we only exchange symbols, which leads to misunderstandings and misinterpretations, as the concepts are not aligned.

As emphasized during an Expert Panel during the Summer Computer Simulation Conference and later captured as a summary in (Tolk 2010), the semiotic triangle can be mapped to M&S. As stated in Hester in Tolk (2010): "Modeling resides on the abstraction level, whereas simulation resides on the implementation level." The conceptualization of the real world referent results in the model, which is equivalent to Ögden's concepts; the implementation of the model results in the simulation, which is equivalent to Ögden's symbols. Furthermore, the interpretation of Ögden is applicable to many interoperability problems as well: While we assume that our simulations represent reality, they actually represent the underlying conceptualization thereof. While current standards focus on data exchange between simulations, this necessary activity only addresses the symbol or syntactic aspect. For meaningful interoperation, the alignment of underlying concepts is necessary as well.

The idea of applying semiotics to M&S is not new. M&S experts use layered models to better understand the various aspects since the beginning. Zeigler, Kim, and Praehofer (2000) propose six layers to describe M&S systems and their application: The *Network Layer* contains the infrastructure including computer and network. The *Execution Layer* comprises the software used to implement the simulation. The *Modeling Layer* captures the formalism for the model behavior. The *Design and Search Layer* supports the design of systems based on architectural constraints. The *Decision Layer* applies the capability to search, select, and execute large model sets in support of what-if analyses. The *Collaboration Layer* allows experts to introduce viewpoints and individual perspectives to achieve the overall goal. Zeigler and Hammonds (2007) associate the semiotic levels as follows: Network and Execution layer define what a system can express, therefore they are associated with the *syntactical* aspects; the Modeling layer expresses the meaning of the syntactic elements, therefor it is associated with the *semantic* aspects; Design and Search, Decision, and Collaboration layer address what the M&S application is used for, therefore they are associated with the *pragmatic* aspects.

The Levels of Conceptual Interoperability Model (LCIM) takes a slightly different view, as it focuses more on the semiotics applicable within distributed and federated simulation systems (Tolk et al. 2008). The LCIM defines a *technical* layer on which infrastructure are addressed, the *syntactical* layer that captures agreements on symbols to be used for the exchange of data, the *semantic* layer that captures the data described by the symbols in form of name and structure associations, the *pragmatic* layer that adds context to the data by grouping them into input and output parameters usable within the distributed and federated event, the *dynamic* layer that makes the reaction of the receiving system in form of state changes transparent, and the *conceptual* layer, that addresses assumption and constraints explicitly. Using the categories recommended by Page et al. (2004) in reaction of an earlier version of the LCIM, these layers can be mapped to the semiotic aspects as follows: The technical layer addresses the aspects of *Integratability*, which contends with the physical/technical realms of connections between systems, which include hardware and firmware, protocols, networks, etc. The syntactical, semantic, and pragmatic layer focus on agreements regarding information exchange using data and address the aspects of *Interoperability*, which contends with the software and implementation details of interoperations; this includes exchange of data elements via interfaces, the use of middleware, mapping to common information exchange models, etc. The dynamic and conceptual layer focus on the use of the data after being exchange as well as the underlying assumptions and constraints and address the aspects of *Composability*, which contends with the alignment of issues on the modeling level. The underlying models are purposeful abstractions of reality used for the conceptualization being implemented by the resulting systems. In other words, integratability focuses of the structure of data resulting in symbols, which is equivalent to the syntactical component of semiotics; interoperability focuses on the interpretation of these symbols as data to be exchanged, which is equivalent to the semantic component of semiotics; and composability focuses on the use and the resulting effect of the data exchange once it happened, which is equivalent to the pragmatic component of semiotics. As addressed in multiple publications, meaningful interoperation requires integratability of infrastructures, interoperability of simulation systems, and composability of models.

Summarizing, syntax focuses on structure of what is (integratability ensures that such structures are exchanged), semantics focuses on the meaning of what is (interoperability ensures data exchange between simulation systems), and pragmatics focuses on the motivation and effects to utilize what is in order to accomplish something (composability ensure that the interpretation and effects is governed by appropriate common assumptions and constraints).

2.3 Mathematics: Model Theory and Algorithmic Information Theory

The sections on ontology, epistemology, and semiotics show the need of understanding what a modeling and simulation application is and what it accomplishes. This section focuses on recent results in the domain of mathematics that support the above mentioned need. As these results have been conducted independent of the former work coming to the same results, they build a strong evidence for their validity.

The first results come from *model theory*. Model theory is a subset of mathematics that focuses on the study of formal languages and their interpretations and is recognized as its own branch since around 1950. The descriptions of these sections are mainly derived from Weiss and D'Mello (1997). Selected practical applications of related ideas for M&S challenges were presented in Tolk et al. (2011a), and some of the definitions used in this section are identical with those already introduced in this earlier work. In essence, model theory applies logic to the evaluation of truth represented using mathematical structures. As computer languages are formal languages, and as simulation systems are programed in computer languages, the results regarding truth representation in formal languages can be applied to consistent representation of truth within computer simulations. As truth regarding the same facts and interpretations need to be consistent within M&S applications, the research findings are of significant importance for understanding interoperability and composability challenges. To understand the research results, some selected definitions of terms are needed: language, model, sentence, and theory, as introduced in Tolk et al. (2011a) based on the definitions given in Weiss and D'Mello (1997).

Definition 1 *A language L is a set consisting of all the logical symbols with perhaps some constant, function and/or relational symbols included.*

Definition 2 *A model (or structure) U for a language L is an ordered pair <A, I> where the universe A is a nonempty set and I is an interpretation function with domain the set of all constant, function and relation symbols of L such that a constant symbol is mapped to a constant, a function symbol is mapped to a symbol and a relation is mapped to a relation.*

Definition 3 *A sentence is an assertion that can be assigned the Boolean value of true or false.*

Definition 4 *If U is a model of L, the theory of U, denoted ThU, is defined to be the set of all sentences of L which are true in U.*

These definitions are closely related to the semiotics and semantics defined earlier. The symbols used in the simulation, which includes all symbols used in the formal language, are captured in the language *L*. Applying syntax to these symbols allows to formulate *sentences*. Using the interpretation function as defined as part of the model, sentences can be evaluated to be true or false, which gives them *meaning*. The theory of a model is the set of all true sentences. Hence, if two simulations need to be consistent, they have to have a consistent representation of truth. It should be pointed out that this requirement does not only exist when independently developed simulation system are federated into a new simulation federation. It is also necessary within each simulation system in which the same simulated entity is represented in several procedures or threats at the same time, potentially using polymorphic representations that need to be kept consistent with each other. The two results of model theory that are directly applicable in support of such challenges are *Robinson Consistency Theorem* and *Łoś Theorem.* Robinson Consistency Theorem simply states that the union of two theories is satisfiable under a model if and only if their intersections are consistent, in other words: there is only one interpretation of truth valid in both models. As it is possible that two theories are using different languages and the resulting sentences are not comparable, Łoś Theorem generalizes the idea of expanding a universe through the Cartesian product and defines filters that allow the comparison in a common equivalent representation. This is only a very small subset of model theory insights but already enough to show that the mathematical foundations of interoperability and composability are already laid and have far reaching implications.

As in the proceeding section we see again that structure and symbols as given in the universe *A* of data describing a simulation system is necessary but insufficient to address logical consistency of the represented theories. In order to decide on questions regarding interoperability and composability, the interpretation function is necessary as well. As already stated by Weiss and D'Mello (1997): the same universe can have many different interpretations resulting in many different theories (or versions of truth).

These findings and observations are also supported by the subset of mathematics called *algorithmic information theory* (Chaitin 1974a, 1974b, 1987) that proposes that a theorem deduced from an axiom system cannot contain more information than the axioms themselves do. To get to this insight, the classical theory of information (Shannon 1948) had to be extended from pure information to algorithms that can produce this information. While Shannon focused on encoding of symbols and syntactical expressions, Chaitin extended these ideas to complete programs. However, the programs are not seen as syntactical expressions but as formal languages that produce sentences. If two theorems produce the same sentences they are equivalent. A program that produces a series of sentences bears therefore the same information as an enumeration of all the sentences it produces. Algorithmic information theory seeks for the shortest and most compact form to produce information, as this is the most efficient form to communicate capabilities. The overall result is that we again need to address ontological structures explaining *what is* as well as epistemological structures addressing *what knowledge is* and what we do with it.

Overall, the recent research results described in these three sections on ontology, semiotics, and mathematics are all pointing to the need to radically change our view on M&S in general and M&S interoperability and composability for M&S applications in particular. M&S can not been seen as a series of activities that are harmonized after the fact. Instead, the tools and methods used must be equivalent language classes under model theory. Furthermore, all evaluated aspects unambiguously show the need to address more than data exchange as it is currently supported by standards. The elusiveness of conceptualizations, as addressed by King and Turnitsa (2008), that makes up the uniqueness of M&S has to captured formally by extending the ideas like proposed in Tolk et al. (2011a). The following section will address the implications for M&S interoperability and composability in more detail.

3 IMPLICATIONS FOR M&S INTEROPERABILITY AND COMPOSABILITY

Analyzing the results of the literature research on related formal approaches presented in section 2, we conclude that a systematic view on interoperability and composability challenges must embrace such formal approaches. Current standard approaches are mainly driven by computational and software engineering perspectives that successfully address the earlier introduced domains of integratability of infrastructures and interoperability of simulation implementations. However, they do not address the special elusiveness of conceptualizations. As stated in Tolk et al. (2011b): *"As we are connecting simulated things we need transparency of what we are simulating, as the real world referent use in other interoperability domains has been replaced in the modeling phase by its representing conceptualization in the M&S interoperability domain."* Using the results of section 2, we can synthesize the implications for M&S interoperability and composability into two subsections that address what needs to be addressed by standards and what should not be addressed by standards. To facilitate following the steps of our synthesis, the main ideas proposed are the following:

- We need to represent ontological as well as epistemological elements of the M&S applications.
- This representation addresses the symbols, the syntax, the interpretation as data elements, and the interpretation of functions and operations. To this end, all semiotic aspects need to be addressed.
- The representation can be expressed using the means of model theory and communicated utilizing the results of algorithmic information theory.

As the ideas following this proposal are now formulated in terms of mathematic formalisms, additional recent research results and recommendations regarding complexity and entropy are applicable as well. The second subsection makes the case that standards aim to minimize entropy to avoid ambiguous interpretations. However, M&S entropy needs to be addressed on the levels of integratability, interoperability, and composability. While minimizing entropy is beneficial on the integratability and interoperability level, it may be counterproductive and a disservice to scientific applicability of M&S when applied to the composability level because the combination of models may lead to emergent behavior, among other issues, needed for gaining insight into phenomena (this is further discussed on the following sub-sections).

3.1 What needs to be addressed by M&S Standards

As discussed in the previous section, several works have pointed to the need for separating syntax, semantics and pragmatics of models and simulations. In M&S interoperability the three aspects need to be aligned in order to have a federation that is working. Based on model theory, M&S models are theory generators which means they have a high degree of entropy at the syntactic level (syntactic entropy) at the sematic level (semantic entropy) and at the pragmatic level (pragmatic entropy). Consequently a federation of interoperating models is also a model which means that it is a theory generator. Therefore, interoperability only adds to the entropy level of the resulting model in an unpredictable manner due to the combination of theories from independently developed models. While this can be seen as positive in many ways, in the prevalent view of interoperability entropy is regarded as a measure of disorder and uncertainty and thus standards are used as a way to bring order and reduce uncertainty by minimizing the entropy. At the syntactic levels and below (integratability), standards such as the ANSI and ASCII codes, standardized data types, formats, libraries and languages, or even standardized protocols (TCP/IP, HTTP) have been very successfully applied to reduce uncertainty of interoperating systems. Based on these successes, there is a belief that it is possible to reduce semantic and pragmatic entropy through standards. Assuming that it is desirable to have standardized syntax, semantics and pragmatics, the following need to be addressed:

- *Standardized languages:* As implied by the result presented by Tolk et al. (2011a), this part is addressed by using a standard language such as the extensible Markup Language (XML), which would guarantee syntactic interoperability with no entropy as a common schema description is used across the federation. There is no uncertainty regarding the symbols, and no room for interpretation. Assuming two systems interoperating in the context of model theoretic terms, the sender could use this language to specify information in the form of a collection of sentences generated from its internal theory. This will guarantee that the sentences in each instance document follow a common syntax. The receiver still has to evaluate the semantics of each sentence with respect to its internal theory, but common name spaces can help to address this issue as well.
- *Standardized theory:* If we apply the results of model theory we conclude that common terms and name spaces are not sufficient, we need to be able to assign consistent truth values to the sentences of the standard language as well, i.e. that we need common theory. For a successful standard, i.e., a standard that avoids inconsistent interpretations of truth among a federation, a common theory that can be used to assign semantics between the sender and the receiver must be specified. We have to be careful that the intersection of the theories abide by Robinson's consistency theorem, otherwise the standard theory introduces a third model that must be aligned with the other two which increases the level of entropy instead of maintaining it. The consequence of having a standardized theory with no entropy within the intersection of the systems means that the introduction of new systems will not have any effect on the entropy of the intersection.
- *Standardized simulation:* Even if sender and receiver are using a common theory, the simulators must be able to generate the theory as well. We understand a simulator as a machine capable of generating a finite state machine realization of the model, i.e., a simulation is the generation of a theory by a simulator. For computer simulation this implies that at the pragmatic level the model must be expressed as a regular language in order to be executable by a Turing machine. The pragmatics of the theory in this case is the number of finite state machine realizations that can be derived from the theories without further simplifications i.e. the number of digital computer implementations of the theory. One of those realizations must be standardized to ensure that there is no minimal during interoperability.

To ensure overall consistency in how we express data, what we mean with these data, and how we use this data, we need standardized languages, standardized theories, and standardized simulations. Picking just one or two of this list is going to increase the entropy within the federation instead of keeping it unchanged, introducing additional degrees of freedom in syntactical, semantic, or pragmatic interpretations of the results produced by the resulting federation.

These observations lead us to ask whether standards are the appropriate way for dealing with interoperability beyond the technical level which we discuss in the next section.

3.2 Constraints for M&S Standards

Terrence Deacon (2007, 2008) wrote the journal articles as contributions to a scientifically adequate theory of semiotic processes. He was convinced that in order to understand what information is and how it is used a theory of information is necessary that can unify the physical, biological, cognitive, and computational uses of the concept. In his work, he also developed the idea that information must be understood on its three semiotic levels. As the authors did in section 2.3 of this paper, Deacon identifies the work of Shannon (1948) as foundational on the symbol level of syntax. He explicitly utilizes the idea of Shannon entropy to introduce the idea of syntactical information. Boltzmann's theory of thermodynamic entropy (Deacon 2007) and Darwin's theory of natural selection (Deacon 2008) are used as examples for semantic and pragmatic information. In his work, Deacon concludes the following: *"What makes something information is not something intrinsic, but something extrinsic to its immediate properties and even its causal history. It is a difference that is interpreted to refer to, or mean, something with respect to some functional consequence."* (Deacon 2007, p. 26) Avoiding increase in the Shannon entropy focuses on well-defined symbols in its most essential form. Avoiding increase in the Boltzmann entropy focuses on interrelated symbols and derivable sentences on the semantic level. Both approaches, however, do not and should not take into account *what* they are talking about, or what the intended use on the pragmatic level will be. Deacon (2008) uses Darwin's theory of natural selection to show how the pragmatic application domain defines the interpretation of exchange data in the context of its use making it information.

A formal proof lies out of the scope of this paper, but the implications for standards are the following: current M&S standards – including the most recent version of IEEE 1516-2010 High Level Architecture (2010) – have the objective to reduce the variety of possible data exchange. They enforce a common world view that has to be shared between contributing M&S applications. They keep syntactic and semantic entropy constant by enforcing a common data exchange that is reduced to a subset of the universe in the model theoretic sense that describes the intersection between two or more participating systems. Current discussions about the need for binary connection standards and on-the-wire protocol in industry show the need for even more rigorous agreements allowing for *plug-and-play* solutions. The perfect solution in this direction is minimizing the entropy on all levels, which means perfect agreement on what information is exchanged in which format for which intent, completely removing the need for any mediation of variations.

Following the argument of Deacon, however, the interpretation of data in the pragmatic realm is where the real information is discovered and used. With the exception of Base Object Models (SISO 2006), none of the current M&S standards addresses the pragmatic level of semiotics at all. This can only work if all simulations implement the same theory consistently, such as using Newtonian physics to model physic-based effects. As long as the application domain is ontologically rooted in such a strict one-world-view assumption – there is only one objective reality and all models are simplifications and abstractions of this same reality and therefore in principal can always be mapped back to each other, known as positivism – this can work out, but as emphasized by Tolk et al. (2011b), this assumption holds only for a small fraction of M&S application domains.

Applying the idea of using standards to minimize the entropy to facilitate interpretation will and cannot work on the pragmatic level, as such an approach would standardize the creative process of developing a model. Standardizing a model also means standardizing the answers it can produce. To be able to apply a model successfully in support of scientific work in its application domain, we need maximal flexibility which results in higher pragmatic entropy.

These observations hint towards a paradigm shift in our view on M&S standards. Instead of reducing the variety of possible solutions, the new paradigm should be based on model theoretic foundations and increase the transparency of models on all semiotic levels but defining the boundaries of interpretation.

3.3 A Possible Paradigm Shift in M&S Standards

Even though we have shown that standards beyond the level of syntactic interoperability are limiting the creativity required for successful use of M&S in support of scientific work, they are still helpful in defining an interoperability space in which M&S systems can be guaranteed minimum interoperability and have the freedom to generate theories supporting creativity within suitable boundaries. The boundaries can be defined by what is *observable*, *reasonable* or *physically possible*.

- The simulated entities have real world references that are bounded by physical limitations. These limits are natural laws that cannot be broken. Boundaries describing these limits fall into the category of boundaries describing what is *physically possible*.
- For many applications, rational behavior of the simulated entities can also be assumed. Although the degree of freedom granted by poor physics is bigger, these entities are furthermore constraint by boundaries describing the *reasonable* limits.
- The most general boundary category is defined by *observable* boundaries that are not driven real-world reference limitations.

In general, these boundaries are associated with 'validity' in M&S, as the simulated system is bounded within the conceptual realm by the same limits as the real world reference within the perceived reality. In addition, model theory shows that a model in M&S must be equivalent to the simulation that generates behaviors and outputs based on rules and axioms defined by the modeler. Consequently, validation in the correspondence sense, i.e. that a simulation is valid with respect to observations of the real world referent, is the ability of the model to generate a real or close enough to a real behavior.

Applying the results presented in this paper, there are two conditions under which this idea of validity and by extension standardization is problematic, namely how to deal with *contradictions* and how to deal with *emergence*:

- In general, the model resulting from federating to simulation systems can generate a behavior and its opposite, which means that the rules and axioms are contradictory.
- In a federation, new behaviors can emerge that are not observable by any of the participation simulation alone.

Contradictions result from a combination of rules and axioms in the case of a composition of models if rules and axioms lead to inconsistent truth representations for the same concept, often captured in both simulations, or as a result of rules and axioms being used beyond the scope of the original simulation system. A bridge, e.g., is not only used in support of traffic, but also to allow for support of the power grid, water support, and more. If network models for the water supply, power supply, and traffic are composed into a federation, it is essential to identify the common resource 'bridge.' If this is not done, contradictions can result in the federation, e.g., the bridge can be destroyed due to a truck accident, but water and power are not affected; in the traffic model, the bridge is destroyed; in the other models, the bridge is intact. In the federation, the bridge is destroyed and intact at the same time. Introducing the bridge into all models solves the contradiction, but also increases the entropy. While in the physical world such problems can be solved, in the psycho-social realm this is not the case. The same is true for medical applications, where potentially contradiction theories and derived methods are applied in practice. Additional examples are given by Tolk et al. (2011b).

Within the current framework and understanding of validity as correspondence to reality, we do not have a way of dealing with this issue. That is unless, the notion of validity is augmented with the requirement for truth consistency, i.e., a model cannot produce something and its opposite. This requires a transparency of simulation systems on all semiotic levels, including the pragmatic level, which is not addressed by any existing interoperability standard.

The philosophical implications of requiring models to be consistent are out of the scope of this paper; however, it means that models have to be developed top down, with not only well-written rules and axioms, but also a way to combine and interpret the rules and axioms. In practice, rules and axioms are embedded within simulations which are a simplified, semi-decidable subset of the model, but the same observations about validity apply here as well. *Standards should make clear whether they allow or disallow contradictions*. If contractions are not allowed, mechanisms should be put in place to either make it explicit or strictly forbid it. In particular in application domains where the use of hypotheses, i.e., reasonable assumptions that have not yet been verified by empirical evaluations, is the rule – such as in social or medical environments – allowing contradictions may be a good thing, as the community simply does not know which hypotheses describes the real world object of interest best.

Positivism – or the notion that all knowledge is evidence based and acquired through empirical observation – seems to be a prevalent view in M&S for historical as well as philosophical reasons. However, in addition to generating observed behaviors, in particular composed models can generate behaviors that are not observable anywhere, at least not in any of the composed models alone. This could mean that the behavior has yet to be observed or there are no tools to observe it, or the conditions for it to be observed have only occurred in the simulation because they only exist in the model. The latter case is either called *emergence* or *invalid,* which begs the question whether emergence is desirable or undesirable in M&S whenever emerging behavior in the simulation has not yet been observed in the real world reference. If we assert that it is desirable, M&S standards should be open to allow for emergence and at the same time guarantee consistency. If we assert that it is undesirable, M&S standards should ensure that interoperability results in a closed and consistent theory. In the later case, M&S applications in domains such as human, social, cultural, and behavior (HSCB) models have to be reduced to closed, consistent theories, which in current applications by scientists within these domains is not the case. *Standards should specify whether they allow or disallow emergence.* If emergence is not allowed, not allowing contradictions apply here as well.

An additional aspect of the interoperability space is provided by observing that in general, contradiction and consistency are competing in the sense that it is impossible to guarantee the consistency of non-trivial models while at the same time guaranteeing their completeness. M&S standards therefore cannot guarantee both consistency and completeness for non-trivial cases. Beyond the technical realm, which is very well defined and needs to be rigid and stable, a choice has to be made as to what constitute the interoperability space.

In order to address this issue, standards need to provide a *standard model* along with a set of axioms and rules. This standard model needs to be derived from the requirements and captures the real world reference attributes, and capabilities of interest in a descriptive form for the simulation. This standard model can be used as the basis for validity, i.e., a model is valid if it can produce the same system as the standard models. The standard model can also serve as the basis for developing models while guaranteeing a level of interoperability and grounding. Finally, this standard model serves at the lower bound of the interoperability space on all semiotic levels.

Technical standards today already fulfill this role at the syntactic level and below. At the semantic level, there are several standard models in several communities that serve as a common reference model. These models are often coupled with documents and best practice recommendations that explain what is reasonable and permitted. At the pragmatic level, in particular in the military domain the accreditation process specifies standard simulations for a domain given a purpose (US Modeling and Coordination Office 2006). While these approaches of best practice are sufficient to spell out what can be done, it does not state what cannot be done or is not permissible.

In addition to the concept of using assertions captured in the standard model, a set of *restrictions* in the form of proof obligations – proofs that certain observations will never be made – must be specified to ensure that certain behaviors are not produced. This set is used to assert that a model does not exceed a threshold of behavior beyond which it is unacceptable under the limit categories introduced earlier in this paper. While the standard model provides the lower boundaries for the allowed variations, the restrictions build the upper boundaries for allowed variations. Similar concepts are usable to furthermore exclude additional intervals by defining lower and upper boundaries of these exception areas.

The modeling space defined by the standard model and the proof obligations can be further defined within each domain and for a given purpose. For instance, the military domain can specify a standard set of models and a standard set of proof obligations for ground infantry training based on manuals, doctrines and experience. This also extends the notion of interoperability from information exchange and use, to the notion of interoperability as a minimum set of information that can be exchanged and used without contradiction and/or emergence.

4 SUMMARY

The work presented in this paper shows the urgent need to rethink current best practices for M&S standards and their application. Current interoperability standards are based on information exchange standards addressing the data to be exchanged between federated simulations without requiring transparency of the sending or receiving simulations regarding how the data was produced or how the data will be used. From a semiotic standpoint, this leaves the pragmatics completely opaque. Also, syntax and semantics are only agreed upon outside of the participating systems, as they are not derived from the internal representations of data and event triggers. This contradicts current findings in the domains of ontology and semiotics that clearly show that all semiotic levels need to be addressed. It also contradicts recent findings of model theory and algorithmic information theory that show the need of model alignment in all phases and over all levels. *Interoperability and composability cannot be engineered into a solution after the fact.* However, simulation systems can still be federated meaningfully, but it requires transparent models to allow deciding whether to simulation systems can be federated. Standards that exclusively focus on information exchange without addressing how the data are used in the receiving system are insufficient to ensure interoperability and composability.

Using model theoretic insights, the paper recommends defining a standard model derived from the requirements that governs the research question to be evaluated by the simulation federation. This standard model builds the lower boundary regarding what shall be produced by the federation. Using additional restrictions that can be motivated by what is physically possible, what is reasonable, or what generally is observable, upper boundaries and areas of exclusions can be defined. In these allowed areas, models may show contradictive behavior or emergence. Standards are needed to govern all these.

The reigning factors to decide what of these is desired is driven by the role the M&S application is applied in. Reynolds (2008) distinguishes two major roles for M&S: Using simulation to solve problems by providing knowledge, and using simulation to gain insight by supporting understanding. In addition, the use of simulation to stimulate testing, in particular hardware-in-the-loop test, shall be considered to exemplify when contradictions, emergence, and restrictions are desirable. Table 1 shows the viewpoint of the authors of this paper.

Table 1: Interoperability Space Characteristics on M&S Roles

	Testing	**Problem Solving**	**Gaining Insight**
Standard Model	*desirable*	*desirable*	*desirable*
Contradictions	*undesirable*	*undesirable*	*desirable*
Emergence	*undesirable*	*desirable*	*desirable*
Restriction	*desirable*	*desirable*	*undesirable*

Lower boundaries as given by the standard model are always desirable, but an upper border may stand in the way of gaining new insight. Contradictions for testing or problem solving are not desirable, but when we gain insight, they may be a necessary first step to find out, where additional research is needed. Emerging behavior – if based on internally valid axioms and rules – can support problem solving and in any case supports gaining new insight, but it is undesirable for testing.

It is important to note that if a standard does not allow contradiction and emergence, the interoperability space is reduced to the standard model with minimal entropy, which makes it a good standard in support of training. On the other hand, a standard that allows all four has high entropy beyond the standard model. Only for the standard model, the entropy remains stable, beyond it, entropy always grows.

None of the current M&S interoperability standards comes close to these requirements, as they were predominantly driven by physics-based modeling requirements under the constraints of positivism. As M&S is increasingly applied in new domains that are not governed by these constraints, a radical change in our view on standards and their role is needed, and – as we are focusing in computer simulations - these new views should be capture formally.

REFERENCES

Bozkurt, I., J. J. Padilla, and A. Sousa-Poza. 2007. "Philosophical Profile of the Individual." In *Proceedings of the 19th IEEE International Engineering Management Conference* (IEMC), 42-48. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chaitin, G. J. 1974a. "Information-Theoretic Computational Complexity," *IEEE Transactions on Information Theory* 20:10–15.

Chaitin, G. J. 1974b. "Information-Theoretic Limitations of Formal Systems," *Journal of the Association for Computing Machinery (ACM)* 21:403–424.

Chaitin, G. J. 1987. *Algorithmic Information Theory*, Cambridge University Press, Cambridge.

Deacon, T. W. 2007. "Shannon-Boltzmann-Darwin: Redefining Information. Part 1," *Cognitive Semiotics* 1:123-148.

Deacon, T. W. 2008. "Shannon-Boltzmann-Darwin: Redefining Information. Part 2," *Cognitive Semiotics* 2:167-194.

Frigg, R., and J. Reiss. 2009. "The philosophy of simulation: hot new issues or same old stew?" *Synthese* 169:593–613.

Hester, P. T., and A. Tolk. 2010. Applying methods of the M&S spectrum for complex systems engineering. In *Proceedings of the 2010 Spring Simulation Multiconference*, Article 75 , 8 pages. San Diego, California: Society for Computer Simulation International.

IEEE 1516-2010: *IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA).* Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

King, R. D., and C. D. Turnitsa. 2008. "The Landscape of Assumptions." In *Proceedings of the Spring Simulation Multiconference*, Edited by H. Rajaei, G. A. Wainer, and M. J. Chinni, 81-88. San Diego, CA: Society for Computer Simulation International.

Ögden, C. K. and I. A. Richards. 1923. *The Meaning of Meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. London, United Kingdom: Routledge & Kegan Paul.

Page, E. H., R. Briggs, and J. A. Tufarolo. 2004. "Toward a Family of Maturity Models for the Simulation Interconnection Problem." In *Proceedings of the Simulation Interoperability Workshop*, Article 04S-SIW-145, 11 pages, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Reynolds, P. F. 2008. "The Role of Modeling and Simulation." In *Principles of Modeling and Simulation: A Multidisciplinary Approach,* edited by J. A. Sokolowski and C. M. Banks, 25-43. John Wiley & Sons.

Shannon, C. E.. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27:379–423, 623–656.

Silver, G. A., O. Hassan, and J.A. Miller. 2007. "From Domain Ontologies to Modeling Ontologies to Executable Simulation Models." In *Proceedings of the 2007 Winter Simulation Conference*, Edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 1108-1117. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Simulation Interoperability Standards Organization (SISO). 2006. *Base Object Model (BOM) Template Specification*. SISO-STD-003-2006. Orlando, FL: Simulation Interoperability Standard Organization

Tolk, A. 2010. "M&S Body of Knowledge (BoK): Progress Report and Look Ahead." *SCS M&S Magazine 2010 No. 4*, 1-5, accessible via http://www.scs.org.

Tolk, A., and J. A. Miller (eds). 2011. *Special Issue: Enhancing Simulation Composability and Interoperability using Conceptual/Semantic/Ontological Models, Journal of Simulation* 5(3):133-245.

Tolk, A., C. D. Turnitsa, and S. Y. Diallo. 2008. "Implied Ontological Representation within the Levels of Conceptual Interoperability Model." *International Journal of Intelligent Decision Technologies* 2(1): 3-19.

Tolk, A., S. Y. Diallo, H. Herencia-Zapana, and J. J. Padilla. 2011a. "Model Theoretic Implications for Agent Languages in Support of Interoperability and Composability," In *Proceedings of the 2011 Winter Simulation Conference,* Edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, 309-320. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Tolk, A., S.Y. Diallo, J.J. Padilla, and C.D. Turnitsa. 2011b. "How is M&S Interoperability different from other Interoperability Domains?" In *Proceedings of Spring Simulation Interoperability Workshop*, 11S-SIW-008, 12-20, Orlando, FL: Simulation Interoperability Standard Organization.

Turnitsa, C. D., J. J. Padilla, and A. Tolk. 2010. "Ontology for Modeling and Simulation," In *Proceedings of the 2010 Winter Simulation Conference,* Edited by B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, and E. Yücesan, 643-651. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

US Modeling and Simulation Coordination Office (2006). *Verification, Validation, and Accreditation (VV&A) Recommended Practices Guide (RPG).* Alexandria, VA, accessible via http://vva.msco.mil.

Weiss, W. and C. D'Mello .1997. *Fundamentals of Model Theory.* Department of Mathematics, University of Toronto, Canada, accessible via http://www.math.toronto.edu/weiss/model_theory.pdf.

Winsberg, E. 2010. *Science in the Age of Computer Simulation.* Chicago: The University of Chicago Press.

Zeigler, B. P., and P. E. Hammonds. 2007. *Model and Simulation-based Data Engineering*. New York, New York: Elsevier Science & Technology Books, Academic Press, 2007.

Zeigler, B. P., T. G. Kim, and H. Praehofer. 2000. *Theory of Modeling and Simulation, 2nd ed.* New York, New York: Academic Press.

AUTHOR BIOGRAPHIES

ANDREAS TOLK is Professor of Engineering Management and Systems Engineering with a joint appointment to Modeling, Simulation, and Visualization Engineering at Old Dominion University. He is also affiliated with the Virginia Modeling Analysis and Simulation Center. He holds a M.S. and Ph.D. in Computer Science from the University of the Federal Armed Forces in Munich, Germany. His email address is atolk@odu.edu.

SAIKOU Y. DIALLO is Research Assistant Professor at the Virginia Modeling, Analysis and Simulation Center at Old Dominion University. He received a B.S. in Computer Engineering, and a M.S and Ph.D. in Modeling and Simulation from Old Dominion University. His email address is sdiallo@odu.edu.

JOSE J. PADILLA is Research Assistant Professor at the Virginia Modeling, Analysis and Simulation Center at Old Dominion University. He received an B.S. in Industrial Engineering from the Universidad Nacional de Colombia, an M.B.A. in International Business from Lynn University, and a Ph.D. in Engineering Management from Old Dominion University. His email address is jpadilla@odu.edu.