A Call to Arms: Standards for Agent-Based Modeling and Simulation

Andrew Collins, Mikel Petty, Daniele Vernon-Bido and Solomon Sherfey (2015)

Journal of Artificial Societies and Social Simulation 18 (3) 12

Abstract

Standards are as old as civilization itself and they are vital to human development. Standards touch almost every part of our lives, from the water we drink to the language used to write this article. A sign of a good standard is one that we do not notice. Good standards exist and so do processes and organizations to create and maintain them. As agent-based modeling and simulation matures as a methodology, a discussion of standards applicable to it becomes increasingly important. Descriptive standards for agent-based models, such as the Overview, Design concepts, and Details protocol and agent-based extensions to the Unified Modeling Language, have already begun to emerge. Software tools for implementing such models, such as Netlogo and Repast Simphony, are increasingly well-known and have the potential to become de facto standards among the wider scientific community for agent-based simulation. Based on the findings of a series of workshops that brought together experts throughout the modeling and simulation community, we argue that agent-based modeling and simulation is no different from the other emerging technical subjects in the sense that standards, both existing and new, may be applicable to it, and that the community should both adopt existing standards that are relevant and exploit the already existing standards processes and organizations to develop new ones.

Keywords:
Agent-Based Modeling and Simulation, Standards, Standardization, Standards Development Organization, ODD, Simulation Methods

Introduction

1.1 Agent-based modeling and simulation (ABMS) has matured as a methodology since Thomas Schelling first introduced his segregation model in the early 1970s (Schelling 1971). It has been applied to problems in defense (Hill et al. 2004; Banks & Sokolowski 2010; Sokolowski et al. 2012; Collins et al. 2014), land use (Polhill et al. 2008), marketing (Rand & Rust 2011), biology (Marino et al. 2008), social sciences (Gangel et al. 2013; Patel et al. 2012), and transportation (Pluchino et al. 2014; Wijermans et al. 2013). The spread has occurred due both to the increasing availability of computer power to the mainstream research and analyst community over the last fifteen years and the usefulness of ABMS as a modeling technique and research tool (Heath & Hill 2010).

1.2 As ABMS is used in a variety of different research and analysis domains, the underlying paradigms in these domains are incrementally being integrated into, or challenged within, ongoing ABMS research. The flexibility of ABMS is evidenced by the range of techniques with which it has been integrated, including social networks (Miller & Page 2007; Seiler et al. 2011), reinforcement learning (Valluri et al. 2009; Collins et al. 2014; Barbosa & Petty 2014a, 2014b), and Latin hypercube sampling (Collins et al. 2013; Marino et al. 2008). However, integration with these different techniques compounds the usual questions of implementation correctness and modeling validity. Too often a cultural preference for fancy computer graphics over well-constructed algorithms and equations has led to the questionable use of visualization rhetoric in simulation outputs, especially commercial packages (Collins & Knowles Ball 2013). Thus there is an emerging motivation for standards in ABMS to combat potential misuse of the approach.

1.3 The majority of the material discussed in this paper was drawn from a series of workshops conducted by the Virginia Modeling, Analysis and Simulation Center (VMASC) on various aspects of M&S standards (Collins et al. 2010a, 2010b; Collins et al. 2011a, 2011b, 2011c). The workshops had over one hundred participants with a wide range of backgrounds and substantial expertise in the development and uses of modeling and simulation standards.

1.4 This paper is organized into six main sections, intended to move from the general to the specific. First, brief background information on the history, breadth, and importance of standards in general is provided. Second, standards developed for modeling and simulation, and the processes used to develop them, are described. Third, some existing standards for ABMS are reviewed. Fourth, an extended example is used to illustrate the applicability of M&S standards to ABMS. Fifth, potential pitfalls and the way forward for developing standards for ABMS are discussed. Finally, the paper is summarized and conclusions are stated.

Standards in General

2.1 Informally, a standard is a way of doing or making something, commonly agreed upon and unambiguously documented that is intended to reduce risk, increase productivity, or enhance interconnectivity. Opinions of standards vary; some claim that standards stifle innovation while others claim that they enable it, e.g., the USB (Universal Serial Bus) interface being commonly quoted as an enabling standard (Collins et al. 2010). One truth about standards is that you cannot get away from them. The water pouring out of the tap in the morning when you brush your teeth is produced to certain standards, as is the cup-holder in your car where you place your coffee. You may not notice the standards around you because, arguably, one of the qualities of a good standard is that the consumer does not need to think about it. In the western standards-driven world, you probably do not worry about the health risks of using contaminated water when brushing your teeth nor do you worry about whether your cup of coffee will fit in your car's cup holder when you purchase it from a drive-thru outlet.

2.2 The use of standards is as old as the spoken word and they have played a fundamental part in human development. Toffler et al. (1981) identified three major waves of human development: Agrarian, Industrial and Information. This list was expanded by Kreckmer (2000), one of the authors of the International Electrotechnical Commission's (IEC) report on standardization as a strategic tool; Kreckmer's list also includes Hunter-gatherer and post-information. Each historic age has certain categories of standards that impact society's development through that age; these are known as successions of standards and are shown in Table 1.

| Age            | Hunter-gatherer | Agrarian | Industrial | Information | Post- |
|----------------|-----------------|----------|------------|-------------|------|

Table 1: Ages of Human Development and their associated standards (IEC 2006)

http://jasss.soc.surrey.ac.uk/18/3/12.html
2.3 Given the long history and clear importance of standards, it is surprising that there are many challenges that still have not been solved within the standards community, e.g., return on investment (ROI). A key observation from the project’s workshops (Collins et al. 2010a, 2010b) was a lack of standardization of many M&S terms and definitions including the word “standards.” A formal definition of a standard, as defined by the US Federal Office of Management and Budget circular A-119 (Office of Management and Budget 1998), is as follows:

   *a. The term ‘standard,’ or ‘technical standard’ as cited in the Act, includes all of the following:
   (1) Common and repeated use of rules, conditions, guidelines or characteristics for products or related processes and production methods, and related management systems practices.
   (2) The definition of terms; classification of components; delineation of procedures; specification of dimensions, materials, performance, designs, or operations; measurement of quality and quantity in describing materials, processes, products, systems, services, or practices; test methods and sampling procedures; or descriptions of fit and measurements of size or strength.

b. The term “standard” does not include the following:
(1) Professional standards of personal conduct.
(2) Institutional codes of ethics.*

2.4 While useful, this definition is certainly not final, and there is dissatisfaction in some who use it (Finkleman 2007). It does not expand upon some of the subtleties of standards, e.g., their status as de facto, de jure, or voluntary. A de jure (formal) standard has been officially sanctioned and certified by a recognized standards body such as the IEEE (Institute of Electrical and Electronics Engineers). A de facto (informal) standard lacks such formal approval but by virtue of its widespread adoption has the benefits of a standard (Morse et al. 2010). A voluntary standard is one that people choose to follow, e.g., fair trade products, and may be formal or informal. It is worth noting that following and compliance to standards is a common practice for practicing engineers.

2.5 Given a working definition of a standard, the next question is who deals with them? There are two broad categories by which the organizations that deal with standards can be categorized. They are commonly referred to as Standards Bodies and Standards Development Organizations (SDOs). Standards Bodies are groups that exercise control over the standards process, which includes managing the acceptance of standardization projects, maintenance of standards over time (to include their withdrawal if necessary), periodic review, certification of standard developers, laboratories, auditors, and other functions, promulgation of acceptable standards development practices, and collaboration and harmonization between national, regional, and international standards efforts. Standards Bodies are usually political entities; if not actually government agencies then Non-Government Organizations (NGOs) assigned the authority and responsibility by governmental decree. In contrast, SDOs are focused on the actual technical work of analyzing the requirements, researching the range of possible variables, parameters, processes and methods, and ultimately creating the technical document that satisfies the needed standardization. While the majority of SDOs are commercial or professional organizations, they may also be government agencies.

2.6 The number and relationship of Standards Bodies and SDOs vary widely between nations. Some can be quite complex; for example, the US structure is illustrated in Figure 1. To give the reader an idea of the extent and scope of standards development activities in the US, they includes some 275 accredited SDOs that are represented by one small icon near the center of the figure, and the Department of Defense (DoD), which is a heavy user of standards and has a $500 billion budget, is represented by a small rectangle in the bottom-left corner of the diagram.

2.7 Similar to the geographic categorization of standards described above, there is a geopolitical hierarchy of Standards Bodies (SBs) that includes NSBs, Regional SBs, and International SBs. Regional and International SBs are composed of representatives of the NSBs that participate, or by government representatives particularly in the case of standards impacting international treaty agreements. Key International SBs and their treaty status include:

- The International Standards Organization – ISO. Non-treaty, founded in 1947. In English, it is easy for us to assume the name ISO is acronymic for the organization, but in fact the name is derived from the Greek word iso, meaning equal.
- The International Electrotechnical Commission – IEC. Non-treaty, founded in 1906.
- The International Telecommunications Union – ITU. Treaty, founded in 1865.
- Codex Alimentarius. Treaty, founded in 1963.
- The Universal Postal Union – UPU. Treaty, founded in 1874.

2.8 All of these International SBs function as voluntary consensus based standards organizations. However, those that are treaty based carry regulatory authority through the sovereign nations participating in the treaty and therefore adhering to the standards they develop. All of these SBOs also function as SDOs, using various means to compose technical committees from their membership and other interested parties to develop their standards.

2.9 Each country (or economy) generally has a single recognized National Standards Body (NSB). Each nation’s NSB plays a uniquely formed role in that nation’s standards environment. They may be the only standards body acting in that nation, with no supporting SDOs, or in large economies they may be solely a coordinating body with hundreds of SDOs in support. Each nation’s NSB is likely their sole representative to the ISO and IEC, and, although they may not be the political entity that represents that nation in treaty organizations such as the ITU or Codex Alimentarius, they are very likely involved in a technical consultant role. A Standards Body may also be an SDO – for example, the British Standards Institution (BSI) is the National Standards Body (NSB) for Great Britain, and is also one of Great Britain’s key SDOs. On the other hand, American National Standards Institute (ANSI) in the US is strictly the NSB – they do not develop any standards at all, although they play an active role in the certification of SDOs to produce them.

2.10 The American National Standards Institute (ANSI) is a non-profit organization with members from both the public and private sectors. Both ISO and ANSI are Voluntary Consensus Standards Organizations (VCOSO). As such, ANSI requires that all associated SDOs follow certain processes for the governance of their standards development including openness of the process; lack of dominance by single interested category, person, or organization; and documented consensus voting (ANSI 2014). However, ANSI approval of a standard does not guarantee the standard’s success.
Standards for Modeling and Simulation

3.1 As an academic discipline, M&S draws its background from a mixture of mathematics, systems engineering, industrial engineering, physics, and computer science. It is relatively new as a discipline, with the serious study of M&S being less than sixty years old, and it has been argued that M&S is still in its theoretical infancy (Tolk et al. 2011). As the discipline matures, new practices, approaches and paradigms will emerge, including the development of new standards within the M&S community. As an area of economic activity, however, M&S has emerged from the twentieth century as a multi-billion dollar industry (Commonwealth of Virginia Secretary of Technology 2014) and as a U.S. National Critical Technology (Forbes 2007).

3.2 The early days of M&S were dominated by bespoke simulation until 1958 when Keith Tocher wrote the first general purpose simulation software while at the United Steel Company in Sheffield (Tocher & Owen 1960). Known as the General Simulation Package, that software soon became the standard tool used for developing simulations, if only for a while. This is first recorded example of simulation built for "reuse" (Hollocks 2006). By reuse, we mean a simulation that can be used for multiple purposes, that is, the simulation is used to answer multiple questions. There are some that do not like this term "reuse" as they argue that simulation should be built for a purpose and, as such, has a modeling intent (Tolk et al. 2011).

3.3 Today an extensive range of standards is available for M&S, including at least 22 individual standards in nine categories that include distributed simulation, conceptual modeling, and simulation development (Petty et al. 2012). That list is certainly not exhaustive. Another summary of current M&S standards can be found in NATO Handbook AMSP-001 (2009). The available M&S standards serve different modeling paradigms and multiple domains, and new standards are under development that will expand their coverage. It seems reasonable to ask why so many M&S standards exist. They are usually claimed to support the reuse of data and models, simplify the integration of models, and reduce development costs. The intent is that if a model or data set is built to a particular standard then it is easier to determine whether it can be used for another purpose and to do so if appropriate.

3.4 For example, consider M&S standards for distributed simulation. In distributed simulation, large simulation systems are assembled from a set of independently executing models running on multiple computational nodes communicating via a computer network. As a distributed simulation system executes the networked models report to each other the attributes (e.g., location) and actions (e.g., firing a weapon) of interest regarding the entities (e.g., tanks and helicopters) they are simulating by exchanging network messages. Standards known as interoperability protocols define the format of the messages, the conditions under which specific messages should be sent, and the proper processing for a received message. Several standard interoperability protocols have been developed over the last 30 years. Three interoperability protocols currently in widespread use are Distributed Interactive Simulation (DIS) (IEEE 2012), High Level Architecture (HLA) (Dahmann et al. 1998; IEEE 2010a, 2010b, 2010c), and Test and Training Enabling Architecture (TENA) (Powell 2010). Although their technical details differ substantially, all three allow models executing concurrently on separate computational nodes to interoperate, i.e., to collaboratively simulate a common scenario (Morse et al. 2006; Tolk 2012; Loper & Turmita 2012; Petty & Gustavson 2012).

3.5 These interoperability protocol standards are widely assumed to reduce development costs. If a model is built to use a particular standard interoperability protocol, then the developers of that model need only to consider that standard; they do not need to consider the communication requirements of all possible models that the model may be expected to interoperate with. The cost saving is in comparison to the development effort that would have been spent developing the model if the standard had not been available. However, quantifying this cost saving is a problematic issue which has not yet been fully resolved by the M&S community (Collins et al. 2012; Oswalt et al. 2012).
3.6 The primary organization that develops M&S standards internationally is the Simulation Interoperability Standards Organization (SiSO), which is based in the US but has strong connections to international organizations, including the Society of Modeling and Simulation International (SCS) and the North Atlantic Treaty Organization (NATO). Two of the previously two mentioned M&S standards, DIS and HLA, were developed by academic/industry teams but are now maintained by SiSO. Both standards benefited from significant international participation in their development processes, and perhaps as a result, both are now widely used both in the US and internationally. SiSO is an SDO as previously defined: SISO-developed standards are often accredited by the Institute of Electrical and Electronics Engineers (IEEE) acting as the SB; this is true of both DIS and HLA. SiSO is a VCSO as defined by the American National Standards Institute (ANSI). Voluntary, consensus based standards development is ANSI's approved method by which a Standards Development Organization (SDO) should be run.

3.7 SiSO is not the only organization developing standards for M&S. For example, the US Department of Defense (DoD), traditionally the world's biggest purchaser of M&S, has developed very large simulation systems. As such, the DoD also has the ability to impose M&S standards on others within its realm and it is worth noting that it has developed its own internal standards, e.g., TENA (Collins et al. 2010b).

3.8 At present, these organizations have not developed any standards that uniquely relate to ABMS. However, some of their standards are broadly applicable to all four the major M&S paradigms: Discrete Event Simulation, Continuous Simulation, Monte Carlo Simulation and ABMS (Law 2014).

3.9 Given the apparent benefits of M&S standards (e.g., cost-saving) and the existence of processes and organizations to develop them (e.g., SiSO) it might seem that the development of useful standards would proceed smoothly. However, "would" is not as the same as "will". Though standards are believed to save time and money they also cost time to produce and some have to provide that time and money. Even if the time and money is available it still takes a period of time to produce a standard and thus there will always be a lag between the latest technology/methodology and its related standards. This would be fine in a relatively stable industry but M&S is a new subject and in a constant state of flux; even the underlying paradigms of M&S are changing (Siebers et al. 2010). Does this mean that standards are not appropriate? No, standards enable M&S practitioners to do things that would otherwise likely be impractical. An example is the very large increase in simulation interoperability enabled by distributed simulation standards. It has even been argued that every simulation in DoD has been influenced by standards (Morse et al. 2010). However, it does mean that any M&S standard is likely to have a limited period of usefulness, and that a standard's life-cycle and its future generational updates need to be considered in its development (Collins et al. 2010b, p. 11); such updates also cost time and money.

3.10 The next questions about M&S standards are 'What should be standardized?' and 'What can be feasibly standardized?' From our workshops, there was a call from the community for more data and verification and validation standards (Collins et al. 2010b, p. 11). Heath et al. (2009) identify verification and validation as major issues for ABMS. Tolk et al. (2011) asserts that standards need to be: valuable (is there a market?), desirable (is there seem to be a problem?), and reasonable (i.e. feasible). Unfortunately, the last two requirements are qualitative in nature thus unappealing to a community, like the M&S community, which tends to be quantitative-thinking in nature.

Standards for Agent-based Modeling and Simulation

4.1 ABMS can be defined as simulating the actions and interactions of autonomous agents where the agents interact with each other and the environment according to rules of behavior. Agents may have the ability to make independent decisions and behave in self-selected and varying ways as determined by their behavior rules, their internal state, and the environment. Agent interactions may include sending messages, exchanging resources, or cooperating to achieve goals. It is an "approach to modelling systems composed of autonomous, interacting agents" (Macal & North 2010), and a modelling paradigm "...with the defining characteristic that entities within the target system to be modelled - and the interactions between them - are explicitly and individually represented in the model ... in contrast to other models where some entities are represented via average properties ..." (Galán et al. 2009). Agent-based models are well suited to model those situations "...characterized by the presence of a number of autonomous entities whose behaviors (actions and interactions) determine (in a non-trivial way) the evolution of the overall system" (Bandini et al. 2009). The agents' rules of behavior and the interactions between the agents may combine to produce emergent behavior, i.e., complex behavioral effects not explicitly encoded in the agents' behavior rules.

4.2 ABMS is an emerging paradigm (no pun intended) and, as such, potential standards for ABMS have begun to be proposed, though not formally governed. Examples include the ODD protocol and AGENT UML, both will be discussed below. There has been recent push towards a community for ABMS (Janssen et al. 2008) but the focus of this community is on communication and not standards development. As with ABMS in particular, M&S in general is still maturing, but it is perhaps farther along, especially in the area of standards. With that in mind, there have been suggestions that standards and standards development processes for M&S should be adopted by the ABMS community (Health et al. 2009):

[a] remedy is that techniques, philosophies and methods need to be adopted from other simulation paradigms, or developed specifically for ABM, and these techniques, philosophies and methods need to be taught to those using ABM such that they can build more effective models.

4.3 Thus techniques, approaches, and standards used by the wider M&S community should not be dismissed out of hand and an example of applicability is given in the next section. However, two relatively new possible standards are unique to ABMS. Neither is a de jure standard, with the approval of an official standards governing organization, nor is either a de facto standard, with broad acceptance throughout the ABMS community. However, both have characteristics of a potential standard, and thereby suggest the relevance of standards to ABMS. These ABMS-specific possible standards are ODD and AGENT UML.

ODD

4.4 Agent-based models (ABMs) allow researchers to explore systems of interaction between autonomous agents. They are used in multiple disciplines to study complex systems comprised of unique, individual entities (Epstein & Axtell 1996; Grimm et al. 1999; Tesfatsion 2002). ABMs can correspond to relatively straightforward systems of interaction or extremely completed ones. Thus the ABMs can themselves become very complex to understand and to analyze. Grimm et al. (2006) attempted to provide clarity to ABMs by standardizing the published ABM descriptions using the ODD (Overview, Design concepts, and Details) protocol. They believed that ABMs were not in greater use because there was difficulty in understanding and duplicating them. Rules, equations, and schedules in ABMs were often not clearly described (Grimm et al. 2008). Grimm (2002) believed that a standard protocol that made the model's description independent of structure, purpose and implementation would benefit all involved. The ODD protocol structures information about an ABM in a standard sequence of seven elements divided into three main sections.

4.5 As its name suggests, the ODD protocol's three main sections are Overview, Design concepts and Details. The overview section is intended to describe the purpose of the model, define the entities, state variables and scale of the model and present the process overview and scheduling of the model (Grimm et al. 2010). This section is designed to communicate clearly the overall objective of the model as well as provide the agents' descriptions, and a description of the spatial and temporal environment. The process overview and scheduling portion is intended to describe the function of each entity, the order in which functions are performed and the process by which state variables are updated. From this section, the reader should have a good understanding of the overall system function.

4.6 The next section of the ODD protocol is the design concepts. The design concepts were originally described by Railsback (2001). Railsback adopted the concepts from the study of complex adaptive systems to help provide a theoretical framework for ABMs. Railsback believed that a set of guidelines would help modelers identify and avoid inappropriate assumptions in modeling. Grimm et al. (2010) believed these guidelines would clarify the communication of the theoretical background and model assumptions. The ODD protocol requires the modeler to describe the model's design for the following seven concepts: emergence, adaptation, fitness, prediction, sensing, interaction, stochasticity, collectives and observation (Grimm et al. 2006).

4.7 The emergent concept describes the occurrence of system level phenomena that is not imposed through behavior rules. Adaptations are the functions that allow agents to change themselves in response to their environment. Fitness functions should describe the means by which an agent's fitness is measured and the methods by which the agent seeks to improve its fitness. The prediction concept is intended to explain how agents estimate the future consequences of their decision.
possibilities. The sensing section of the design concepts details the state variables that an agent is assumed to know— that is, the known state variables of the agent and of the “neighboring” agents and environment. Interaction tells the type of interactions that occur between agents and the environment. Stochasticity describes the randomness and the rationale for the randomness of the model and collectives are used to define any group structures that exist. The final design concept is observation. This concept is intended to explain the data collection process used for the analysis of the model. Railsback (2001) notes that ABMs are “unstable if the patterns of individual behavior cannot be observed.”

4.8 In the ODD protocol, the overview provides a high level review of the model’s functionality and elements. The design concepts described the theoretical framework and assumptions of the model. The final section of the ODD protocol describes the model details. The model initialization parameters, including the initial environment and agent creation and states should be described. The inputs to the model need to be detailed. This should include describing any static data sets as well as describing any methods used to generate data sets. Lastly, the sub-models should be described. The sub-models should contain all the rules, underlying assumptions and parameter values used in the model. This detail should allow for the complete reproduction of the model.

4.9 At first glance, the ODD protocol can seem tedious and redundant. The purpose of the model is typically stated more than once in an article using the protocol and the sub-models are listed in both the scheduling and the detail sections (Grimm et al. 2010). However, the standard enforces a discipline on a traditionally ad hoc design. The standardization of communication appears to provide greater rigor in the model development (Grimm et al. 2010).

AGENT UML

4.10 The Unified Modeling Language (UML) is used by the computer science community to diagrammatically represent (“model”) a software program or a conceptual model. This allows developers to communicate ideas succinctly to each other in a standard way (Chonoles & Scharf 2011). Given that most agent-based simulations beyond Schelling’s work have been constructed using computer software, using UML to describe the agent based model seems quite reasonable. The UML 2.0 standard has a collection of 14 different box-arrow diagrams that are used to explain different aspects of the model, e.g., class diagrams represent the different classes used in object-oriented programming and the relationship between them. There have been two attempts to apply UML to ABM, that is the AGENT UML (Bauer et al. 2001) and a direct application (Bersini 2012).

4.11 Bauer et al. (2001) introduced AGENT UML which is an extension of UML 1.4, which was the current version of UML at the time. They argued that UML did not work for ABMs because agents are adaptive and work in communities (Bauer 2002). They proposed extensions for standard UML diagrams: class, sequence, package and templates. Ultimately the extensions to UML focused on the use of logical operators to take into account the complexities of ABMS.

4.12 Independently of Bauer’s work, Bersini (2012) proposed a method for using UML 2.0 for representing ABMS. Due to differences between UML 1.4 and UML 2.0, e.g., the lack of a standard syntax for timing or communication protocols, Bersini was able to propose a methodology of using UML 2.0 without the need for extensions. Bersini’s approach focuses on the four diagrams: class, sequence, state and activity. Note that last two diagrams are different from Bauer et al. (2001), which can be accounted for by the difference between the two versions. Thus due to improvement in the UML standard, ABMS can now be represented within UML without the need for an ABMS-specific syntax.

4.13 The application of ODD and UML standards to ABMS was derived from a desire, of the ABMS community, to better express the inner workings of an agent-based model. Though UML is an existing standard, extensively used by the computer science and engineering community, it requires the user to be comfortable with block-arrow diagrams. The ODD protocol, on the other hand, requires the user to be able to give an eloquent prose description of their model, something that might escape those from a technical inclination. Thus, no conclusions are given here about which standard is the most appropriate. We believe that it is up to the ABMS community to discuss the future usage of such standards.

4.14 Both the ODD and UML focus strictly on the communication of model’s technical details. This is not the only area where standards could be applied to ABMS. For example, standards could be derived for sampling the parameter space in agent-based simulation runs. Sampling techniques like Latin hypercube sampling have been applied to ABMS but require assumptions about the parameters distribution (Collins et al. 2013; Marino et al. 2008). Standards could be derived to determine what is considered acceptable assumptions for sampling by the ABMS community. We have purposely tried to avoid a general discussion for potential future ABMS standards in this paper because we believe it is distracting from our main discussion point: standards should be derived in formal and inclusive way.

An Example of the Applicability of M&S Standards to ABMS

5.1 To illustrate the potential applicability of M&S standards to agent-based models, we offer an extended example and argument in this section. First, a type of entity-based combat models known as semi-automated forces, which are widely used for military training and analysis applications, are described. Then, although semi-automated forces are not customarily categorized as agent-based models, definitions of the latter from the literature are invoked to argue that the former do in fact belong in that category. Finally, the significant relevance of standards, especially distributed simulation interoperability protocol standards such as the previously mentioned DIS and HLA, to the utility of semi-automated forces systems is explained.

5.2 Military simulations often include simulated entities (such as tanks, aircraft, or individual humans) which are generated and controlled by computer software rather than by human crews or operators for each entity. (This is a familiar feature of many computer games as well.) In the military context, the entity-based combat models that generate and control such entities are known as semi-automated forces (SAF) systems, where “automated” applies because software generates much of the entities’ behavior automatically and “semi” applies because the system is monitored and optionally controlled or overridden by a human operator.

5.3 In a military training application, SAF systems are often used to generate opponents against which human trainees engage in virtual battles. Doing so with a SAF system is preferable to having additional human crews in simulators control the hostile forces because SAF systems are both less expensive, as they reduce the need for a large number of simulators not available for the trainees, and more flexible, in that they can be configured to use the tactical doctrine of a particular adversary more readily than retraining human opponents. SAF systems can also generate friendly forces, allowing a small group of trainees to practice teamwork within a large friendly force. In non-training simulation applications, such as analysis (e.g., testing a revised tactical doctrine or assessing the effect of an enhanced weapon), SAF systems typically are used for adding all of the entities involved in the simulation, allowing the analysis scenarios to be executed repeatedly to support statistical analysis without exhausting human operators.

5.4 The entities generated and controlled by the SAF system exist in a battlefield that is a simulated subset of the real world, so the physical events and phenomena on the battlefield must be modeled within the SAF system. For example, if a SAF vehicle is moving, its acceleration, deceleration, and turn rates on different terrain types must be modeled. Combat interaction need to be modeled in accordance with the physics of weapon and armor performance characteristics.

5.5 SAF systems use specialized algorithms to generate the behavior of the entities they control that allows those entities to react autonomously to the battlefield situation as represented in the simulation (Perry 2009). The most widely used approach to behavior generation in SAF systems has been finite state machines (FSMs) (Smith & Petty 1992; Calder et al. 1993). In an FSM implementation an entity’s behavior is decomposed into a finite set of behavior patterns, or states. Associated with each state is an implementation of that state’s behavior pattern in the underlying programming language. Dynamic conditions within the simulation may trigger transitions from one state to another and thus from one behavior pattern to another. The behavior must be both behaviorally realistic, in that it appears to be similar to human behavior in the same situation, and doctrinally consistent, in that the actions of the SAF-controlled entities should be consistent with doctrine of the entities the SAF is simulating.

5.6 SAF systems provide an interface that allows a human operator to monitor and control the SAF entities’ behavior. Figure 2 shows an example of a typical SAF operator interface. (The image is from the ModSAF system (Ceranowicz 1994). ModSAF has been replaced in military applications by a newer SAF system known as OneSAF (Parsons et al. 2005). We use a screen image from the older system to avoid any concerns with sensitive information.) The operator may input high level plans that are
executed in detail by the SAF system, initiate automatic entity behavior, or manually override software-generated behavior. SAF system interfaces provide a map of the battlefield that shows the battlefield terrain and the simulated entities on it. In the figure, there are three companies of Red entities visible as groups of small icons in the northeast, north central, and west central areas of the map, and one company of Blue entities in the southeast area. The Red entities are all executing a general tactical action known as a “Hasty Attack” and the Blue entities are executing a different tactical action known as “Hasty Occupy Position”. These actions were selected by the operator. The SAF software automatically generates in real time specific autonomous movement and combat behavior for each entity that is consistent with the tactical action, considers the terrain, and responds to the presence and actions of friendly and enemy entities.

Figure 2. Example SAF operator interface.

5.7 We argue that SAF systems satisfy our definition of agent-based models, given at the start of the Standards for Agent-based Modeling and Simulation section, in all important aspects. SAF systems normally model combat forces as individual entities or agents (e.g., tanks or helicopters), not as aggregate units (e.g., battalions). The SAF entities’ behaviors are largely autonomous, in that they are generated by the SAF system software, and independent, in that each SAF entity moves and fires individually based on its own state and its perceived state of the environment. The SAF entities interact with the environment, in that the battlefield terrain affects entity actions including movement and inter-visibility, and the entities interact with other in various ways including formation movement, cooperative tactics, refueling and rearming, and of course direct fire. These interactions all follow the entities’ rules of behavior. Finally, there is a very real sense in which combat, both in reality and as modeled in SAF systems, is the archetypical example of emergence, with the individual entities’ actions dynamically determined by situation-specific and transient
battlefield conditions, and the overall battle outcomes wholly dependent on the cumulative effect of the entities’ interactions.

5.8 While SAF systems can be executed in a standalone mode, they are often part of a distributed simulation system (Petty & Gustavson 2012). In a distributed simulation system, the SAF system may be connected via a computer network to human-suited simulators, other SAF systems, and other types of models and support nodes. The connection allows, for example, human trainees in simulators to interact with SAF entities in a common virtual battlefield. The messages sent and received between the nodes of the distributed simulation system are defined using standardized distributed simulation interoperability protocols. The interoperability protocols most often used by SAF systems are the DIS and HLA standards (Petty & Gustavson 2012).

5.9 A large part of the utility of SAF systems depends on their ability to interoperate in this manner, and all major SAF systems have this capability. Clearly such standards are relevant to SAF systems. Given our previous argument that SAF systems are agent-based models, then it follows that modeling and simulation standards have the potential to be relevant to agent-based models, at least some types of them. Indeed, although distributed simulation interoperability protocol standards are relevant to agent-based models by the argument just made, they are not the only relevant modeling and simulation standards. The Base Object Model standard for conceptual modeling, the Distributed Simulation Engineering and Execution Process standard for systems engineering of large simulation systems, and the Verification, Validation, and Accreditation Recommended Practices Guide for testing models are all potentially useful for the development and testing of agent-based models.

Discussion and the Way Ahead

6.1 The discussion on standards has so far highlighted their importance to humanity and the existing processes and organizations that exist to develop and govern them. As ABMS matures a technique, the need to consider standards for it increases. There already exist procedures that the ABMS community can use, e.g., VCSO, and there already exist organizations that can be leveraged, i.e., SISO. However, the ABMS community may be slow to embrace standardization due to perceived difficulties. Some of these perceived difficulties are discussed here.

Uniqueness of ABMS

6.2 Is ABMS unique is some way? Is M&S unique in some way? It has been argued that M&S is its own subject area, as opposed to a subset of some other academic subject such as systems engineering, in part because it has its own unresolved problems, e.g., comparability (Weisel et al. 2003; Davis & Anderson 2004; King 2009) and M&S interoperability (Tolk et al. 2011). In regards to M&S standards, beyond dealing with the unique discipline of M&S, are their development and governance really any different from the thousands of other standards development organization? The generally accepted view from our workshops was “probably not.” M&S standards are comparable to software standards, hardware standards or business process standards. The advantage of this statement is that the relevant research and lessons learnt from other standards fields can be applied to M&S standards and there are a lot of other standards to choose from. The problems faced by non-M&S standard development organizations are similar to those faced by the M&S community. Support for this statement can be found throughout the literature, for example when Tolman discusses building and construction modeling standards (Tolman 1999); he concludes that:

"ISO is not the optimum organization to steer the pre-standardization process and there is not even consensus among the researchers that are carrying out the efforts. As there is no strong management commitment and no funding, it is not realistic to expect that STEP [the construction modeling standards organization] will solve the industry's problems."

6.3 For a reader familiar with the M&S standards development, it is easy to see that Tolman's statement could have been found in many a report on M&S standards by simply substituting out the organizations names for those relating to the M&S industry. For those involved in ABMS, this statement is possibly the face of things to come.

6.4 All methodologies, subjects and techniques are unique in some way and many of them have their own standards. Thus it is not immediately obvious why ABMS is so unique that it would need a different approach to developing standards other than VCSO approach advocated by ANSI. Indeed, we have shown an example of a class of agent-based models, semi-automated forces, where M&S standards developed using that approach has been highly successful. Determining what would make ABMS unique enough for VCSO not to apply is difficult because there does not exist a common definition of ABMS! For example, consider two definitions of ABMS by North and Macal (2007) and Gilbert (2007). Though this is the common that derive the standards and it will be the ABMS community that decides what standards are necessary. This does not mean that existing M&S standards are not applicable to ABMS, e.g., DIS and HLA.

6.5 As ABMS matures so will its accepted definition by the community, and that definition will become one of its first de facto standards. Thus standards are unavoidable; we are arguing that they should be developed in a logical and well-tested manner, through the successful practices used by VCSOs, such as SISO, operating within the formal guidelines of standards accrediting organizations, such as the IEEE. However, whether this is accepted by the community is a different matter; we would adapt the famous Shakespearian quote to scientific subjects and suggest that "Nothing is so common-place as a scientific community's belief that the problems of their subject are so unique that standard development processes do not apply to them."

6.6 Though we have argued that ABMS is not unique enough to merit its own means of standards development, it is still unique and, as such, might require its own standards that fit the needs and goals of the ABMS community. We reject the notion that ABMS is just a subset of Discrete Event Simulation (Law 2014). Though this might be true at the implementation level of agent-based simulation, it is not true for the purpose of ABMS which includes finding emergent behavior and modeling complex adaptive systems (Miller & Page 2007). Standards are not static and they will evolve, with their own life cycles, as the community grows and matures (Morse 2010). It is the community that derives the standards and it will be the ABMS community that decides what standards are necessary. This does not mean that existing M&S standards are not applicable to ABMS, e.g., DIS and HLA.

Organizational Misbehavior

6.7 Organizational misbehavior, in the context of M&S standards, is defined by Collins et al. (2011) as "behavior by individuals and organizations that slows, de-rails, or otherwise negatively affects the standards development process, whether this is purposely or unintentionally done." Though the effects of organizational misbehavior are well documented, Collins et al. (2011a) documented several anecdotal examples which were given anonymously at the aforementioned workshops: persistent obstructionism within the standards development process, malicious compliance to the standard and sloppy implementation of the standard. The reasons behind such misbehavior depend on motivation of the obstructing person or organization; for example, if they wished to delay the standards development then they might raise specious objections to the standard during the development process (persistent obstructionism). Some misbehavior might occur for no malicious reasons; it may just be that the organizations involved in the development of the standard are at cross-purposes. More discussion on the management consideration of M&S standards can be found in Henninger et al. (2009), which discusses M&S standards development and implementation in the U.S. Department of Defense.

6.8 The examples given at our workshops imply that organizational misbehavior is everywhere but there is very little literature on organizational misbehavior within an M&S standards context. Hollenbach does give a history of the rise and fall of the High Level Architecture (HLA) standard (Hollenbach 2009). Collins et al. (2011a) suggests that organizational misbehavior can be analyzed and, perhaps understood by using game theory. Gardener and Moffat (2008) provide an application demonstration of game theory to this type of issue with their example relating to defense acquisition.

The wrong standards
The mere fact that a standard has been developed does not mean it is well suited for its intended purposes. In some cases, it might be simply wrong. For example, the temperature of the human body was found by the German physician Carl Wunderlich to be 98.6°F (37°C) during the mid-nineteenth century. This standard was used until the 1990s when it was actually shown to be 98.2°F by Mackowiak et al. (1992). In fact, it was shown not to be constant at all and would vary both over the course of the day and over a person's lifetime, typically going down with age (Gomolin et al. 2005). The consequence of an incorrect standard was that patients were misdiagnosed with fevers for more than one hundred years and antibiotics where sometimes given to people who did not necessary need them. The possibility of error suggests that standards should not be static; rather, they should evolve as new information is discovered and new approaches are developed. For this to happen an active standards development community must exist and continue to exist after the standard has been initially defined.

Consequences of not having standards

In contrast to our discussion on standard themselves, the long-term consequences of not having standards is now considered. Some have argued that communication, or interoperability, is at the heart of all M&S standards (Collins et al. 2010a) and hence communication, relating to standards, is the focus of our discussion here.

Difficulties of communication: The ODD protocol has allowed developers ABMS a common format with which to describe the creations to the wider community and UML allows the developers to do this for the more technically minded. These two standards were developed in response to a need to present an agent-based model in a way that was easily accessible to the community and highlights its points of interests as defined by the community. We make no judgment on whether either standard was successful or not at achieving better communication amongst the ABMS community.

Anything will (and probably has) been called ABMS: The lack of consensus on the definition of ABMS means that the academic literature has too many examples of papers describing models which claim to have used ABMS but bear little resemblance to the criteria highlighted by any of the current collection of definitions. It is already difficult to convince people that emergent behavior is useful and scientifically sound; it is going to be harder if people are exposed to "agent-based models" that have no relation to emergent behavior or other key ideas of ABMS like heterogeneity, adaptiveness, etc. (North & Macal 2007). To combat this issue, standards teaching material could be used for ABMS like those advocated by Macal and North (2013). However, as already mentioned there is no current agreement on the definition of ABMS, and so the community has a long way to go before any "standard" educational material on ABMS is prevalent in higher education establishments.

ABMS is regularly confused with multi-agent modeling and simulation: Multi-agent modeling is a useful computer science technique for investigating the interaction of different software agents. The key distinction is that ABMS is focused on modeling real-world systems with agents that are often abstractions of real-world entities, whereas multi-agent modeling includes any computational system consisting of interacting agents, including those where the agents are software agents with no correspondence to real-world entities. Multi-agent simulation has also been applied to simulation applications of game theory, where the "multi" really means "two." Multi-agent systems tend to focus on removing unwanted behaviors and encouraging desirable ones; agent-based modelers want to observe unexpected emergent behavior to better understand the real-world system that has been modeled. The danger of the confusion of the two approaches is that it might be assumed that advances in one cross over to the other. Appropriate standards may help to distinguish the two.

Summary and Conclusions

Standards are all around us, from the size of our coffee cups to the format of this article. They allow us to function and connect in our highly sophisticated technological and socially complex world. Standards also allow people to communicate ideas in a mutually understandable manner. From a commercial standpoint, standards have enabled manufacturers and software developers to produce products more quickly and at less cost. However, there are reasons to fear standards; they can reduce freedom of choice and there are those that would profit from them, e.g., patenting a standard so that all users are required to pay a usage fee.

The focus of this paper was the motivations and mechanisms for the future development of standards for ABMS. To that end, background information on standards in general and their relationship to the development of human society was provided, and M&S standards, including the organizations that develop them and the current accepted processes for doing so, were discussed in some detail. To emphasize the need to follow current standard development methodologies, the paper highlighted a few of the potential pitfalls for standards development that the ABMS community might face in the future and the consequences of not having standards.

The authors, as well as others, e.g., Macal and North (2013), believe that ABMS is qualitatively different from traditional modeling methods such as discrete event simulation, and has the potential to give researchers insight beyond those previously available. However, as the ABMS research community grows it is likely to become increasingly difficult to disseminate methods and approaches without some form of standards. As ABMS matures as a technique and becomes more widespread, the need for standardization will grow. Both formal (de jure) and informal (de facto) standards have the potential to be useful for ABMS. The authors also believe that the development of ABMS standards is inevitable and that some of the issues listed earlier will be faced by the ABMS community eventually.

Fortuitously, ABMS is not the first technical community with a need to develop standards, and as a result, there are existing development and governance processes and organizations to draw from. We argue that the ABMS should embrace the already existing standards procedures, such as those outlined by American National Standards (ANSI), and/or appropriate organizations, like the Simulation Interoperability Standards Organization (SISO). Both of which have already faced a lot of the issues highlighted here and have experience with dealing with them. (None of the authors of this paper have any direct professional connection to SISO.)

However, standards cannot be developed in a vacuum and their development requires both resources and willpower. This article contributes to the discussion on standards in ABMS. It is hoped that, though this discussion, both the ABMS and the M&S standards communities will consider serious engagement into ABMS standards development.

Acknowledgements

The paper would not have been possible without the insightful contributions of the hundreds of participants that took part in the series of workshops that some of discussion is drawn from. The workshops were all held under a non-attributable understanding (aka Chatham House Rules) which has been respected. The workshops were supported by the Department of Defense (DOD) Modeling and Simulation Coordination Office under contract N6133910-C-0033, whom also acted as our technical customer. Finally, the authors acknowledge the anonymous referees, whose well-considered comments helped to substantially improve this article.

References

AMERICAN NATIONAL STANDARDS INSTITUTE (ANSI) (2008). Standards and Conformity Assessment Bodies of the United States (2006-07-21 ed.). New York, NY: ANSI.

AMERICAN NATIONAL STANDARDS INSTITUTE (ANSI) (2014). ANSI Essential Requirements: Due process requirements for American National Standards (January 2014 ed.). New York, NY: ANSI.

BANDINI, S., Manzoni, S. & Vizzari, G. (2009). Agent Based Modeling and Simulation: An Informatics Perspective. Journal of Artificial Societies and Social Simulation, 12(4), 4. http://jas.soc.surrey.ac.uk/12/4/4.html.

BANKS, C. M. & Sokolowski, J. A. (2010). Modeling the Niger Delta Insurgency. The Social Science Journal, 47(2), 271–293. [doi:10.1016/j.soscij.2009.11.005]
of the 2012 AlaSim International Modeling and Simulation Conference.

PETTY, M. D. & Gustavson, P. (2012). Combat Modeling with the High Level Architecture and Base Object Models. In: Tolk, A. (ed.), Engineering Principles of Combat Modeling and Distributed Simulation, Hoboken NJ: John Wiley & Sons, 413–448. [doi:10.1002/9781118180310.ch19]

PLUCHINO, A., Garofalo, C., Inturri, G., Rapisarda, A. & Ignaccolo, M. (2014). Agent-Based Simulation of Pedestrian Behaviour in Closed Spaces: A Museum Case Study. Journal of Artificial Societies and Social Simulation, 17(1), 16: http://jasss.soc.surrey.ac.uk/17/1/16.html [doi:10.18564/jasss.2338]

POLHILL, J. G., Parker, D., Brown, D. & Grimm, V. (2008). Using the ODD Protocol for Describing Three Agent-Based Social Simulation Models of Land-Use Change. Journal of Artificial Societies and Social Simulation. 1(2), 3: http://jasss.soc.surrey.ac.uk/11/2/3.html

POWELL, E. T. (2010). The Benefits of TENA for Distributed Testing and Training. Proceedings of the 2010 Interservice/Industry Training, Simulation, and Education Conference, Orlando FL, November 29-December 2 2010 , 1157–1206.

RAILSBACK, S. F. (2001). Concepts from complex adaptive systems as a framework for individual-based modelling. Ecological modelling, 139(1), 47–62. [doi:10.1016/S0304-3800(01)00228-9]

RAND, W. & Rust, R. T. (2011). Agent-based modeling in marketing: Guidelines for rigor. International Journal of Research in Marketing, 28(3), September 2011, 181–193. [doi:10.1016/j.ijresmar.2011.04.002]

SCHELLING, T. C. (1971). Dynamic Models of Segregation. Journal of Mathematical Sociology, 2(1), 143–186. [doi:10.1080/0022250X.1971.9989794]

SEILER, P. O., Macal, C. M., Garnett, J., Buxton, D. & Pidd, M. (2010). Discrete-event simulation is dead, long live agent-based simulation! Journal of Simulation, 4(3), 204–210.

SIEBERS, P. O., Macal, C. M., Garnett, J., Buxton, D. & Pidd, M. (2010). Discrete-event simulation is dead, long live agent-based simulation! Journal of Simulation, 4(3), 204–210.

SMITH, S. H. & Petty, M. D. (1992). Controlling Autonomous Behavior in Real-Time Simulation, Proceedings of the Southeastern Simulation Conference 1992, Pensacola FL, October 22-23 1992, pp. 27–40.

SOKOLOWSKI, J. A., Banks, C. M. & Morrow, B. (2012). Using an agent-based model to explore troop surge strategy. The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 9(2), 173–186. [doi:10.1177/1548512911401739]

TESFATSION, L. (2002). Agent-based computational economics: Growing economies from the bottom up. Artificial life, 8(1), 55–82. [doi:10.1162/106454602753694765]

TOCHER, K. D. & Owen, D. G. (1960). The automatic programming of simulations. Proceedings of the Second International Conference on Operational Research, Banbury J & Maitland J (eds), English Universities Press: UK, pp 58–60.

TOFFLER, A., Longul, W. & Forbes, H. (1981). The third wave. New York, NY: Bantam books.

TOLK, A., Balci, O., Diallo, S. Y., Fishwick, P. A., Hu, X., Loper, M., Petty, M.D., Reynolds, P. F., Sarjoughian, H. & Zeigler, B. P. (2011). Towards a methodological approach to identify future M&S standard needs. Proceedings of the 2011 Winter Simulation Conference, pp. 2975-2992. [doi:10.1109/WSC.2011.6147999]

TOLK, A. (2012). Standards for Distributed Simulation. In: Tolk, A. (ed.), Engineering Principles of Combat Modeling and Distributed Simulation, Hoboken NJ: John Wiley & Sons, 209–242. [doi:10.1002/9781118180310.ch12]

TOLMAN, F. P. (1999). Product modeling standards for the building and construction industry: past, present and future. Automation in Construction, 8(3), 227–235. [doi:10.1016/S0926-5805(98)00073-9]

VALLURI, A., North, M. J. & Macal, C. M. (2009). Reinforcement Learning in Supply Chains. International Journal of Neural Systems, 19(05), 331–344. [doi:10.1142/S0129065709002063]

WEISEL, E. W., Mielke, R. R. & Petty, M. D. (2003). Validity of Models and Classes of Models in Semantic Composability.Proceedings of the Fall 2003 Simulation Interoperability Workshop, Orlando FL, September 14–19 2003, 526–536.

WIJERMANS, N., Jorna, R., Jager, W., van Vliet, T. & Adang, O. (2013). CROSS: Modelling Crowd Behaviour with Social-Cognitive Agents. Journal of Artificial Societies and Social Simulation, 16(4), 1: http://jasss.soc.surrey.ac.uk/16/4/1.html