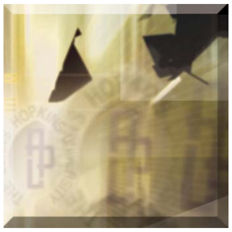


Modeling and Simulation Verification and Validation Challenges

Dale K. Pace

Modeling and simulation play increasingly important roles in modern life. They contribute to our understanding of how things function and are essential to the effective and efficient design, evaluation, and operation of new products and systems. Modeling and simulation results provide vital information for decisions and actions in many areas of business and government. Verification and validation (V&V) are processes that help to ensure that models and simulations are correct and reliable. Although significant advances in V&V have occurred in the past 15 years, significant challenges remain that impede the full potential of modeling and simulation made possible by advances in computers and software. This article identifies major modeling and simulation V&V challenges and indicates how they are being addressed.

INTRODUCTION

From the earliest days of computer modeling and simulation in the 1950s, those who create the models and simulations and those who use them have been concerned about model and simulation correctness and the level of credibility that should be attributed to their results.[1,2] However, only in the last decade or two has there been major emphasis on formal approaches to ensure both model and simulation correctness and credibility.[3] Reports by the Government Accounting Office[4] and others[5] drew attention to modeling and simulation limitations, especially to validity and credibility issues.

As software engineering matured as a discipline, the need for formal, consistent, and repeatable processes in the production of quality software was continually demonstrated. Widespread acceptance of the Software Engineering Institute's Capability Maturity Model (CMM) and its follow-on (CMMI)[6] to judge the maturity of an organization's software processes illustrates recognition of this need. A similar need for advances in modeling and simulation was also recognized. Significant advances have resulted from increased use of the Unified Modeling Language (UML)[7] and other descriptive paradigms facilitating formal consistency in model development, especially those using formal methods.[8] Unfortunately, yardsticks for measuring organizational modeling and simulation maturity comparable to CMM/CMMI for software do not exist except in the arena of distributed simulation specification where compliance with the Distributed Interactive Simulation (DIS)[9] protocol or the High Level Architecture (HLA)[10] provides a reliable indication of interoperability with other simulations. The importance of comparable advances in verification, validation, and accreditation (VV&A) also became clear. Starting in the early 1990s and continuing since, government agencies and professional societies significantly increased their emphasis on more formal V&V/VV&A for software and for models and simulations by establishing and upgrading their policies and guidance related to such.[11] The boxed insert notes some of the Laboratory's contributions to model and simulation VV&A.

During the past 15 years, the availability of modeling and simulation V&V literature has increased substantially. Many books on software V&V/Independent V&V (IV&V) are now available that basically satisfy the need for technical information related to modeling and simulation verification. Many are not aware of the differences between connotations for V&V of software and V&V for modeling and simulation. Some definitions for software V&V describe both verification and validation in terms of specified requirements,[12] which can cause all software V&V to be what is addressed in model and simulation verification (that the developer's intent was achieved). Model and simulation validation goes a step further to determine whether the simulation can support intended use acceptably. A general book on simulation validation was published a decade ago,[13] a specialized book on V&V for computational science and engineering applications was published in 1998,[14] and the proceedings of the Foundations '02 V&V Workshop provided documentation of the current state of V&V practice.[15] A number of modeling and simulation textbooks now have V&V/VV&A chapters, and scores of V&V-related papers, articles, and reports are published each year.

This article begins with a brief review of what modeling and simulation VV&A is. It then identifies several challenges currently facing modeling and simulation V&V and examines how they are being addressed. Implications for modeling and simulation utility from these challenges are also discussed.

SELECTED VV&A CONTRIBUTIONS BY APL PERSONNEL

Community Leadership

- VV&A Technical Director for Defense Modeling and Simulation Office (DMSO); Co-chair/Chair of DoD VV&A Technical Working Group (TWG) and NATO VV&A groups (1997–present)
- Chair of VV&A Group/Forum for Distributed Interactive Simulation (DIS) and Simulation Interoperability Workshop (SIW) (1993–present)
- Co-chair of Tomahawk Simulation Management Board overseeing VV&A of Tomahawk simulations (1983–present)
- Leadership roles for Standard Missile and modeling and simulation VV&A
- Co-chair of major VV&A workshops: Simulation Validation 1999 (SIMVAL99) sponsored by the Military Operations Research Society (MORS) and Society for Computer Simulation (SCS); Foundations '02 sponsored by 28 government, industrial, and professional society organizations

Publications and Education

- Scores of VV&A conference papers, journal articles, and book chapters
- VV&A tutorials and short courses for government, industry, and professional societies
- VV&A policy, guidance, and template development assistance to military services and defense agencies
- Lead/participating role in development of professional society V&V guidance

VV&A Reviews

- Lead role for APL models and simulations
- Lead role for models and simulations elsewhere within the defense community

WHAT IS VV&A?

There are a variety of formal definitions for VV&A terms. Some are authoritative and widely used, such as those in DoD directives and instructions[16] or in professional society standards or guides.[17] Others may pertain only to the particular documents containing them. In general, the connotations of the VV&A definitions are as follows.

Verification

Did I build the thing right? Have the model and simulation been built so that they fully satisfy the developer's intent (as indicated in specifications)? Verification has two aspects: design (all specifications and nothing else are included in the model or simulation design) and implementation (all specifications and nothing else are included in the model or simulation as built).

Validation

Did I build the right thing? Will the model or simulation be able to adequately support its intended use? Is its fidelity appropriate for that? Validation has two aspects: conceptual validation (when the anticipated fidelity of the model or simulation conceptual model is assessed) and results validation (when results from the implemented model or simulation are compared with an appropriate referent to demonstrate that the model or simulation can in fact support the intended use).

There have been many paradigms of the relationships among V&V activities and model or simulation development and what is represented in the model or simulation. One of the earliest and most influential was the "Sargent Circle" developed in the 1970s by Dr. Robert Sargent of Syracuse University,[18] a major figure in simulation validation for the past three decades. Figure 1 is an evolution of that paradigm developed by Sargent

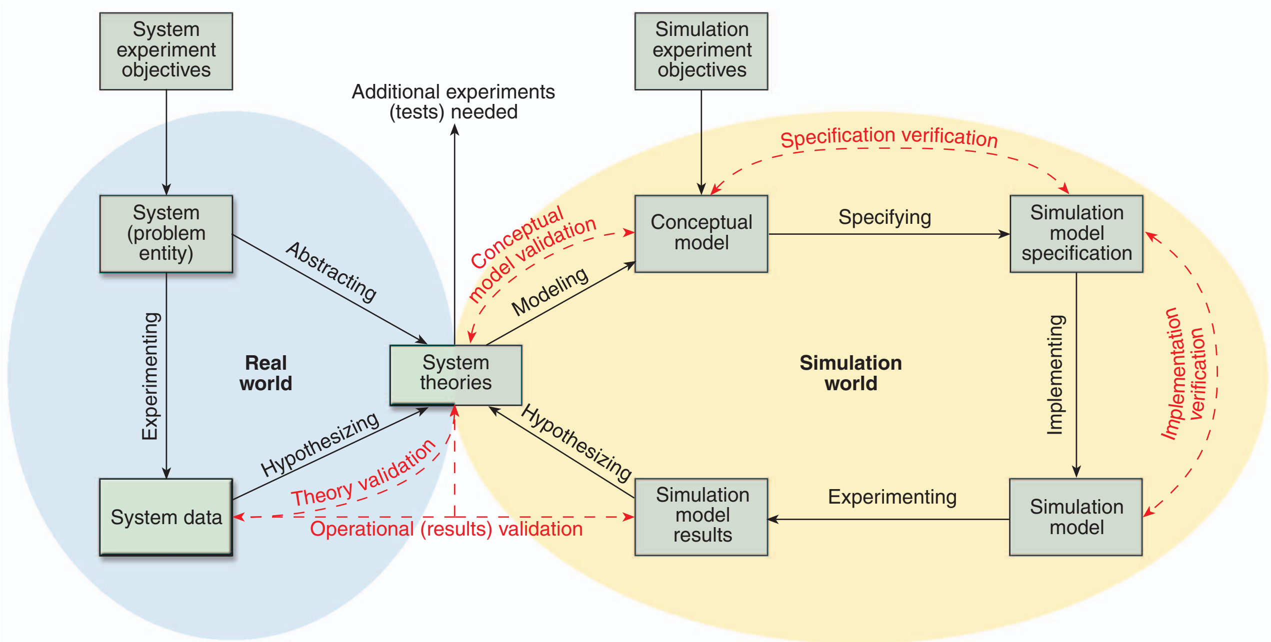


Figure 1. Real-world and simulation-world relationships in developing system theories and simulation models with verification and validation. Experiment objectives should be derived from validated requirements. Validation is always relative to objectives/requirements/intended use. The dotted red lines imply comparison, assessment, or evaluation. (Diagram developed and copyrighted by Dr. R. G. Sargent, Syracuse University, Jan 2001; reprinted with permission.)

in 2001 at this author's request to help the American Society of Mechanical Engineers (ASME) committee develop a V&V guide for computational solid mechanics. This evolution of the Sargent Circle not only shows major V&V activities, but also clearly indicates similar elements in the experimental data (referent) side of VV&A.[19]

The modeling and simulation literature has largely neglected the referent, in the past failing to provide guidance about how to select or describe the information used as the standard for comparison in validation assessment. A recent study chartered by the Defense Modeling and Simulation Office (DMSO) VV&A Technical Director suggests that "the referent is the best or most appropriate codified body of information available that describes characteristics and behavior of the reality represented in the simulation from the perspective of validation assessment for intended use of the simulation" and provides guidance about referent identification, selection, and description.[20] Sometimes the "best" information is too costly, would take too long to obtain, or has some other impediment, so "appropriate" information that has adequate fidelity for the intended use of the model or simulation serves as the referent.

Accreditation

Should it be used? Accreditation is a management decision that may include schedule and other considerations as well as technical V&V information. Authority for the accreditation decision is normally vested in those responsible for consequences from the use of results from the model or simulation. Often the accreditation decision is based on whether the model or simulation can support a subset of its requirements called "acceptability criteria," i.e., capabilities that must be demonstrated for a particular application. Figure 2 illustrates relationships among the requirements, acceptability criteria, V&V activities, and other information used in an accreditation decision.

Credibility

Should results from the model or simulation be believed? Credibility reflects how anyone exposed to model or simulation results should view those results. Typically,

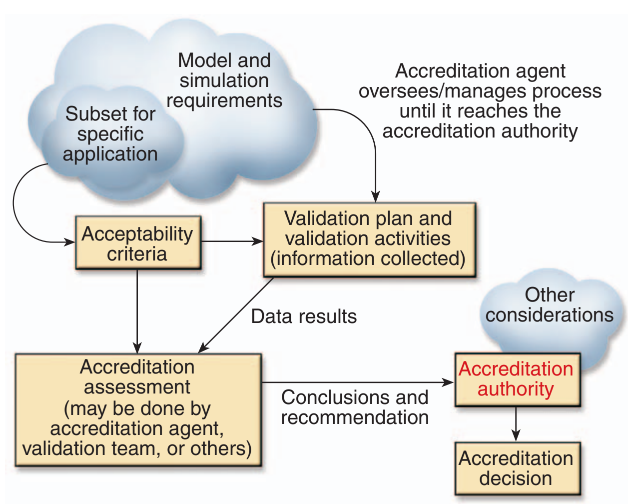


Figure 2. Accreditation in a nutshell.

model or simulation credibility is a function of available V&V information, accreditation status, the reputation of those who developed or use the model or simulation, and its history of use. Sometimes viewer bias or prejudice for or against the model or simulation is also a factor.

V&V CHALLENGES

In October 2002, 198 people from Belgium, Canada, France, Germany, the United Kingdom, and the United States met at APL for the *Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, better known as Foundations '02, which was sponsored by 28 organizations from the United States and abroad (10 government organizations, 8 academic institutions, 6 professional societies concerned about modeling and simulation, and 4 from industry; Fig. 3). The boxed insert identifies the workshop sponsors. The Foundations '02 proceedings[15] describe current modeling and simulation V&V practice at the college textbook level.

The following general conclusions about current modeling and simulation V&V are taken from the Foundations '02 Executive Summary.

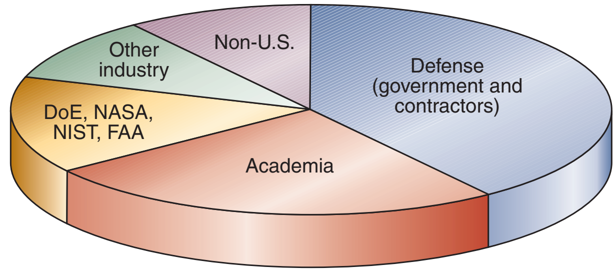


Figure 3. The 198 attendees of Foundations '02 came from the United States, the United Kingdom, Germany, France, Canada, and Belgium.

- The primary motivation for modeling and simulation V&V is risk reduction, i.e., to ensure that the simulation can support its user/developer objectives acceptably. This provides the primary benefit in cost-benefit concerns about V&V, which is the core issue in the question of how much V&V is needed.
- Effective communication is a problem because of continuing differences in the details about terminology, concepts, and V&V paradigms among various modeling and simulation communities; excessive use of acronyms makes it difficult to communicate easily across community boundaries.
- Advances in modeling and simulation framework/theory can enhance V&V capabilities and is essential for increasing automated V&V techniques.
- Limitations in items required for effective V&V (required data and detailed characterization of associated uncertainties and errors, simulation/software artifacts, etc.) must be addressed, with many of the management processes for coping with them being common in many areas of simulation application.
- Cost and resource requirements for modeling and simulation V&V are not as well understood as they need to be because meaningful information about such is not widely shared within modeling and simulation communities, and much more information about cost and resource requirements needs to be collected and made available to facilitate development of more reliable estimation processes.

FOUNDATIONS '02 SPONSORSHIP AND PARTICIPANTS

AEgis Technology Group, Inc.
American Society of Mechanical Engineers Applied Mechanics Division (ASME/AMD)
Arizona Center of Integrative Modeling and Simulation (ACIMS)
Association for Computing Machinery Transactions on Modeling and Computer Simulation (ACM TOMACS)
Boeing Phantom Works
CentER, Tilburg University (The Netherlands)
Clemson University
Defense Modeling and Simulation Office (DMSO; main initiating sponsor)
Federal Aviation Administration (FAA) Flight Standards Service
Gesellschaft für Informatik (Bonn, Germany)
Illgen Simulation Technologies
Innovative Management Concepts, Inc. (IMC)
JHU/APL (facility provider)
Joint Accreditation Support Activity (JASA)
Joint Army, Navy, NASA, Air Force Interagency Propulsion Committee (JANNAF) Modeling and Simulation (M&S) Subcommittee (initiating sponsor)
McLeod Institute of Simulation Science, California State University (CSU), Chico
Modeling and Simulation Information Analysis Center (MSIAC)
NASA Ames Research Center
National Institute of Standards and Technology (NIST)
National Training Systems Association (NTSA; hosting sponsor)
Office of Naval Research (ONR)
Shodor Education Foundation, Inc.
Simulation Interoperability Standards Organization (SISO)
Society for Modeling and Simulation International (SCS)
Survivability/Vulnerability Information Analysis Center (SURVIAC)
U.K. Ministry of Defense (MoD) Synthetic Environment Coordination Office
U.S. Association for Computational Mechanics (USACM)
University of Central Florida Institute for Simulation and Training (UCF/IST)

- Areas of modeling and simulation V&V need to employ more formal (repeatable and rigorous) methods to facilitate better judgments about appropriateness of simulation capabilities for intended uses.

The modeling and simulation V&V community is faced with two very different kinds of challenges. One set relates to modeling and simulation management (or implementation): *how to do what we know how to do* in a proper manner consistently. The other challenges have a research flavor: *areas that we need to understand better in order to find viable technical solutions*. This article identifies and discusses challenges of both varieties.

Management Challenges

Foundations '02 identified three management (implementation) challenges: (1) *qualitative assessment*, (2) appropriate and effective use of *formal assessment* processes, and (3) model and simulation/V&V *costs/resources* (accounting, estimation, benefit). The challenge is how to ensure that "best practices" are employed where they exist and are pertinent.

1. Qualitative assessment. Qualitative assessment involves human judgment in assessment: "peer review," "subject matter expert" (SME) evaluation, face validation, etc. Often people involved in qualitative assessments are selected and perform their assessments without appropriate credentials or formal processes. Methods exist that, if used, can increase qualitative assessment objectivity and consistency.
2. Formal assessment. Formal assessment, whether statistical in nature or following some other rigorous mathematical approach, can be difficult to employ fully. The management challenge is to develop appropriate "lightweight" variants of the processes that can be more easily used in modeling and simulation V&V to enhance the quality of formal assessments.
3. Costs/resources. Correct estimation of resources needed is a primary challenge in any modeling and simulation application. Information available for reliable estimation of modeling and simulation V&V costs and needed resources is inadequate. The management challenge is to collect and organize appropriate cost and resource information (from case studies and other sources), and make it available to the modeling and simulation communities so that robust methods for model and simulation/V&V cost/resource estimation can be developed.

Research Challenges

Foundations '02 identified four areas of research challenges: (1) *inference*, (2) coping with *adaptation*, (3) *aggregation*, and (4) *human involvement/representation*.

1. Inference. Data availability to support assessment of simulation "predictions" is a fundamental problem, especially for the test and evaluation community on the operational side and the experimental community on the science side. Comparison of simulation results with the available data can be described statistically, and data-simulation result relationships can be specified in terms of accuracy, error, resolution, etc., for the region of the application domain for which data exist; *but there are currently no scientifically rigorous methods for making inferences about simulation results (predictions) elsewhere in the application domain (i.e., in those regions where data do not exist)*.
2. Adaptation. Advances in technology have led to a new genre of computational programming, sometimes termed complex adaptive programming. Techniques employed in adaptive programs include artificial intelligence, expert systems, genetic algorithms, fuzzy logic, machine learning, etc. As adaptive processes become more capable and more widely incorporated in modeling and simulation, the V&V challenge is clear: the model and simulation structure and parameters can differ from one run/iteration to the next as the program adapts, and this presents fundamental challenges to the prediction and assessment of performance. *No scientifically rigorous methods currently exist to ensure that future modeling and simulation performance involving adaptive programming will be as good as or better than past performance*.
3. Aggregation. Elements and interactions of a simulation can be represented in varying levels of detail. As simulations become more complex, especially in distributed simulations which may use more than one level of resolution for the same kind of element or interaction, better methods for determining the potential impact on simulation results from such variation in levels of detail are required to minimize potential misuse of simulation results. *Present theory and assessment processes related to this topic are embryonic*.
4. Human involvement/representation. Representation of human behavior in simulations is widely recognized as being critical; the complexity of representing the variety of human behaviors in an automated way that appropriately reflects impacts of the simulated situation on human decision making and performance is a major challenge. The critical stumbling block is uncertainty about influences of factors and processes involved for many kinds of simulation applications. Although better understanding exists about simulation V&V when people are involved for education/training purposes or when used to represent human behavior in the simulated situation, many significant research issues remain concerning interactions among simulation characteristics, the people involved, and appropriate simulation uses.

CURRENT PROGRESS IN OVERCOMING V&V CHALLENGES

Foundations '02 identified the foregoing modeling and simulation V&V challenges. This section discusses progress being made to overcome those challenges.

1. Qualitative Assessment

Many areas, from most medical diagnoses to knowledge engineering and surveys, mainly depend on qualitative assessment. In these various disciplines, a great deal of effort has been expended to ensure that the qualitative assessments are generally credible and, to a reasonable extent, repeatable (i.e., the assessment would be the same regardless of which practitioner made the assessment). Qualitative assessment in modeling and simulation V&V does not enjoy that kind of reputation: "SMEs commonly provide unstructured, vague, and incomplete evaluations."[21] Such judgments are anecdotal since there are no large, careful studies of how qualitative V&V assessments are performed; however, a variety of modeling and simulation assessments by many different sources all leave the impression that there is much room for improvements in this area and that the credibility of qualitative assessments has not improved much with time.

Foundations '02 provides suggestions that would enable modeling and simulation V&V qualitative assessments to have the same level of credibility and repeatability found in qualitative assessments in other arenas.[22] Many SME V&V assessments fail to take even the most fundamental steps to enhance the credibility of their assessments, steps as basic as being explicit about what was reviewed and what were the logical and factual bases of the assessment. It is little wonder that so much doubt exists about the correctness of simulation results when V&V depends so much on qualitative assessments.

Although there are some experiments with qualitative assessments (such as one to see if the use of formal survey techniques can enable SME predictions to be a more credible surrogate referent in the absence of data[23]) and some efforts to improve qualitative assessments for individual simulations, no widespread efforts are under way to better these assessments in modeling and simulation V&V. This is particularly unfortunate since the new Joint Capabilities Integration and Development System (JCIDS) process for defense system acquisition places increased emphasis on qualitative assessment ("warfighter judgment") in an important area of DoD use of model and simulation results.[24]

2. Use of Formal Assessment Processes

Formal methods have been used to a great degree in computer hardware design, but the conclusion of a formal methods expert in 1998 was pessimistic: "[L]imited progress has been made in applying formal methods to software,… [and] the use of formal methods in practical software development is rare."[25] The two areas in which formal methods have been used most are security and safety applications. Foundations '02 noted significant progress, particularly with lightweight formal methods, and pointed out that formal methods "may be the only approach capable of demonstrating the absence of undesirable system behavior."[26] However, formal methods are not used as much as they could be in modeling and simulation, in part because those who develop and design models and simulations lack appropriate math training. In addition, there is no indication that this situation is likely to change in the near future. Most of the research, publications, and work in and with formal methods continue to occur outside the modeling and simulation communities.

3. Modeling and Simulation/V&V Costs/Resources (Accounting, Estimation, Benefit)

Many efforts have been made to estimate the benefits of modeling and simulation[27] and V&V. All of these efforts are limited not only by a lack of information about V&V costs but also by a lack of information about modeling and simulation resource costs[28] and the lack of widely accepted ways to measure modeling and simulation benefits.[27,29] Despite the increasing reliance on models and simulations in strategic planning, conceptual development, and system design and manufacturing as well as more effective and creative ways to use existing systems, it is unclear that adequate information is currently being collected about modeling and simulation costs or their V&V to develop reliable methods for estimating required resources. Foundations '02 has a useful summary of current methods for estimating V&V costs.[29] This author is unaware of any widespread effort to collect information about resource expenditures for modeling and simulation, or their V&V, in ways that facilitate or enable accumulation of data upon which useful resource estimation techniques might be based.

4. Inference

The computational science and engineering community has been in the forefront of the inference issue, with a major emphasis on the importance of being able to describe and quantify uncertainty, both in the model or simulation and in the experimental data (or other referent materials such as benchmark cases used in model calibration).[30] Progress in quantifying uncertainty (whether for experimental data, observations and theory, or model and simulation results) is necessary for more meaningful statements about the relationship between simulation results and referents used in validation assessments as well as for progress in inference.

Foundations '02 summarized the issues involved with inference, as illustrated by Fig. 4.[31] The challenge

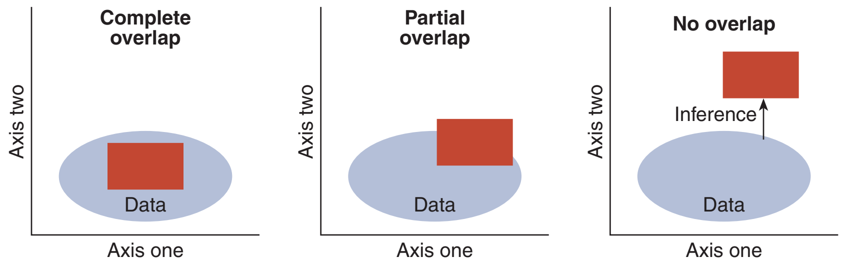


Figure 4. Possible relationships between validation referent data (ovals) and application domain (squares).

is having adequate information to make predictions in regions that differ from the ones in which data are available.

In the past few years, many modeling and simulation communities have become much more aware of the importance of being explicit about uncertainties associated with models, simulations, and the information used as the standard in validation and accreditation assessments. It is now more likely that such uncertainties will be discussed explicitly in quantitative terms instead of largely being ignored, as often happened in the past. Unfortunately, the basic inference issue—how to estimate the accuracy of simulation predictions with scientific certainty in regions beyond those containing validation data—remains unsolved.

5. Coping with Adaptation

Models and simulations that employ adaptive programming change during operation. For some, the changes are relatively minor, e.g., a few weighting factors might change as the program is trained. For others, the changes are major, affecting the structure of the program as well as individual processes within the program. While much of the research and analysis of complex adaptive systems occurs outside modeling and simulation communities, these communities are extensively involved in this area, especially in the arena of agent-based simulations. An agent may have its own world view, be capable of autonomous behavior, communicate and cooperate with other agents, and exhibit intelligent behavior. Foundations '02 identified the basic V&V issues associated with adaptive programs.[32]

During the past few years, a great deal of experience has been gained with models and simulations employing adaptive programs; however, no significant progress has been made in V&V methods related to such models and simulations. We still have no way to prove that future simulation results will be as good as past results, but we blithely expect them to be at least that good.

6. Aggregation

Major models and simulations normally aggregate some representational aspects, either to make the model easier to use (e.g., less input data would be required) or to allow the simulation to execute in a reasonable time. Typically, such aggregation is not in the primary areas of interest. However, some models and simulations aggregate aspects that are of primary interest, and that is the focus here. More thinking about aggregation in this way has been done at the RAND Corporation than elsewhere within the United States, at least in the arena of defense-related models and simulation. The B6 paper of Foundations '02, later expanded by its authors and published as a RAND report,[33] captures most of the validation issues related to aggregation and presents a helpful paradigm for relating results from aggregated simulations to various referents. The authors use multiresolution, multiperspective modeling and exploratory analysis to validate models for specific contexts.

Progress in aggregation V&V is hindered by the lack of comprehensive modeling and simulation theory that is widely accepted and used throughout modeling and simulation communities. Such a comprehensive theoretical context is essential for evaluating abstractions used in aggregation to determine if those abstractions are compatible and consistent. It takes a comprehensive and coherent theoretical modeling framework, such as the discrete event system specification (DEVS) developed and promulgated by Dr. Bernard Zeigler of Arizona University and his associates, to allow logical and mathematically defensible assessments of abstraction appropriateness when aggregation is employed in a simulation, as illustrated by DEVS extensions into HLA applications and elsewhere.[34] Unfortunately, most models and simulations are not developed in such formal environments. The object-oriented approach employed for many models and simulations does not have mechanisms to ensure coherence among the objects developed and their abstraction.

Today, as in the past, validation assessments of aggregation appropriateness in a single simulation, or among various simulations whose results are used together in an analysis, depend on the skill and adroitness of the analysts involved, and there is no substantial evidence that analysts today do this better than those of the past. Consequently, credibility issues continue for simulations using aggregation.

7. Human Involvement/Representation

During the past decade, improving the capability to represent human behavior in models and simulations has been consistently emphasized and pursued within the DoD and elsewhere. Advances in the computerized representation of human performance coupled with advances in computer speed have made better human behavior representation (HBR) possible, but there are still significant limitations in contemporary HBR, as illustrated in an abstract for a recent presentation on the subject: "Human behavior representation (HBR) is an elusive, yet critical goal for many in the simulation community. Requirement specifications related to HBR often exceed current capabilities. There exist a number of tools, techniques and frameworks to model and simulate HBR, but to work they must be constrained and so do not generalize well."[35] Some are candid about the need to keep HBR fidelity requirements as low as possible in order to make model or simulation development affordable.[36] The synopsis of current HBR capabilities (relative to model and simulation validation) from Foundations '02 is still applicable.[37]

Less formal attention has been paid to the validation of the appropriateness of humans used in simulations to represent human performance, and this author has not seen indications that validation processes in this regard have improved in recent years. This area still requires significant attention in order for simulation credibility to improve.

THE BOTTOM LINE

Substantial advances have been made in modeling and simulation verification; this article has focused mainly on validation and has not addressed verification in as much detail. The ability to prevent or detect and correct faults when the capabilities of computer-aided software engineering tools, other kinds of automation, and formal methods are employed is substantially greater than was the case a decade ago, but unfortunately, these capabilities are not routinely used to their fullest potential. Foundations '02 provides a convenient synopsis of current modeling and simulation verification technology.[38]

Foundations '02 divided current modeling and simulation challenges into two groups as we noted earlier: management or implementation and research. As indicated above, little progress seems to have been made in the past 2 years in the management or implementation arena. In this, the modeling and simulation culture has been as resistant to changes as other cultures, and improvements are likely to be very slow (at least in the opinion of this author). Only if serious sanctions are placed on those choosing to continue to use past "business as usual" ways instead of ensuring that current best practices are used consistently in all aspects of modeling and simulation V&V does this author see hope for significant improvement. Serious sanctions are likely to become common only when modeling and simulation results are held to liability standards similar to those for hardware. Administrative guidance, not even when stated in standards or DoD directives and other formal documents, seems unable to change modeling and simulation behaviors in ways that are needed to improve the credibility of modeling and simulation results.

During recent years, three of the four research-oriented challenges (inference, adaptation, and HBR) have received substantial attention, and progress has been made, even though no major validation breakthroughs have occurred in these areas. The fourth area of challenge (aggregation) has received far less attention. In part, this is because research in aggregation depends on progress in the areas of general modeling and simulation theory as well as in processes of abstraction. Modeling and simulation communities seem willing to leave these more as art than science. We do not want to have to grapple with formal relationships among the traditional abstraction mechanisms (i.e., classification, generalization, and attribution) and their relationship to context in a limited domain, such as database structures,[39] or in the broader modeling and simulation arena.

While research in these four areas is not coordinated or focused by any single guiding principle, progress is being made. However, there seems to be relatively little communication across community boundaries in regard to these research areas, and no formal effort to synthesize insights from them all as they might pertain to model and simulation validation. Even those organizations with general perspectives and responsibilities to integrate available information from all sources seem to be stuck with a stovepipe mentality and focused in narrow areas.[40] Some, such as the Modeling and Simulation Technology Research Initiative (MaSTRi) at the University of Virginia, seem to appreciate that these areas are interrelated. MaSTRi's focus is on "the solution of critical challenges that have inhibited or prevented the use of modeling and simulation technology in otherwise practical settings. Critical challenges include multi-resolution modeling, interoperability, visualization, behavioral modeling, security, confidence assessment, visualization in architectural engineering environments and integration of modeling and simulation (M&S) into training and education."[41] Even where such vision exists, as it does in many organizations, not much synthesizing or synergetic use of modeling and simulation research is evident, at least not as far as it pertains to overcoming the four research areas of validation challenge addressed here.

CONCLUSION

This article has reviewed the seven V&V challenge areas for modeling and simulation identified by Foundations '02 and has presented the author's impression of progress during the 2 years since the workshop. Some may feel that the view expressed here is too jaundiced. The author wishes he could present a more positive picture, but this assessment is based on what he finds in the current literature, hears at various conferences and workshops, observes happening in various modeling and simulation endeavors, and learns from personal communication with many in various modeling and simulation communities in the United States and elsewhere.

It is important to remember that current modeling and simulation has far greater capacities than in the past, and that modeling and simulation V&V is better than it used to be, better in the sense that it is more likely to be done formally and with serious effort to use good practices. However, there is still much room for improvement, as indicated in this article.

The growing reliance on modeling and simulation results to guide system design and operational philosophies increases the importance of making those results acceptable under all circumstances and in knowing clearly where the limits are on their appropriateness.

REFERENCES AND NOTES

[1]Nance, R. E., and Sargent, R. G., "Perspectives on the Evolution of Simulation," *Oper. Res.* **50**(1), 161–174 (Jan/Feb 2002).

[2]Balci, O., and Sargent, R. G., "Bibliography on Validation of Simulation Models," *Newsletter of TIMS College on Simulation and Gaming* **4**(2), 11–15 (1980).

[3]Topcu, O., *Review of Verification and Validation Methods in Simulation*, Technical Memorandum 2003-055, Defense R&D Canada-Atlantic, Dartmouth, NS, Canada (Apr 2003).

[4]For example, *DOD Simulations: Improved Assessment Procedures Would Increase the Credibility of Results*, GAO/PEMD-88-3, Government Accounting Office (1987).

[5]For example, Roth, P. F., Gass, S. I., and Lemoine, A. J., "Some Considerations for Improving Federal Modeling," in *Proc. 10th Winter Simulation Conf.* **1**, IEEE Press, Piscataway, NJ, pp. 213–218 (1978).

[6]Software Engineering Institute (SEI) Capability Maturity Models, available at http://www.sei.cmu.edu/cmm/cmms/cmms.html; and Capability Maturity Modeling Integration, available at http://www.sei.cmu.edu/cmmi/.

[7]Unified Modeling Language, available at http://www.uml.org/.

[8]Formal Methods Repositories, available at http://vl.fmnet.info/repositories/.

[9]IEEE 1278.1–4 for Distributed Interactive Simulation (1995–1998, with 2002 revisions).

[10]IEEE 1516/1516.1–3 for Modeling and Simulation (M&S) High Level Architecture (HLA) (2000 to 2003).

[11]No comprehensive listing of such V&V/IV&V/VV&A policy/guidance is available. The briefings and papers of the T3 Session of the Foundations '02 V&V Workshop (proceedings available from the SCS or linked to the DMSO VV&A Web page) identify most of DoD's model and simulation V&V/VV&A policy and guidance as well as those from several professional societies (AIAA, ASME, IEEE, etc.). Not included in these are V&V guidance from NASA (which established a software IV&V center in 1993), V&V guidance from the National Institute of Standards and Technology (NIST) from the 1990s, the substantive V&V efforts begun in the late 1990s as part of the Department of Energy's Accelerated Strategic Computing Initiative (ASCI) program, software validation guidance from the Federal Drug Administration (FDA), or V&V guidance from a number of other government, industrial, and professional organizations.

[12]For example, the description of software validation from Soft Solutions International: "Software validation is essentially a design verification function and includes all of the verification and testing activities conducted throughout the software life cycle," available at http://www.ssi-ltd.com/services/software-validation-ethos-definitions.asp.

[13]Knepell, P. L., and Arangno, D. C., *Simulation Validation: A Confidence Assessment Methodology*, IEEE Computer Society Press, Los Alamitos, CA (1993).

[14]Roache, P. J., *Verification and Validation in Computational Science and Engineering*, Hermosa Publishers, Albuquerque, NM (1998).

[15]Pace, D. K. (ed.), *V&V State of the Art: Proc. Foundations '02, a Workshop on Model and Simulation Verification and Validation for the 21st Century*, Laurel, MD, CD-ROM, published by The Society for Modeling and Simulation (2002); see the VV&A page of the DMSO Web site, http://www.dmso.mil/public/transition/vva/foundations.

[16]DoD Modeling and Simulation Master Plan 5000.59-P, Under Secretary of Defense for Acquisition and Technology, Washington, DC (Oct 1995); DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A), Instruction No. 5000.61, Office of the Director of Defense Research and Engineering, Defense Modeling and Simulation Office, Washington, DC (23 May 2003).

[17]For example, *Software Verification and Validation Plan*, IEEE Standard 1012-1998 (1998), or AIAA *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, AIAA-G-077-1998, Reston, VA (1998).

[18]The Sargent Circle was promulgated by the Society for Computer Simulation in Schlesinger, S., "Terminology for Model Credibility," *Simulation* **32**(3), 103–104 (1979).

[19]The evolved Sargent Circle was first published in Pace, D. K., "Simulation Conceptual Model Role in Determining Compatibility of Candidate Simulations for a HLA Federation," in *Proc. Spring 2001 Simulation Interoperability Workshop (SIW)*, paper 01S-SIW-024, available at http://www.sisostds.org/doclib/doclib.cfm?SISO_FID_5912.

[20]*The Referent Study Final Report* (May 2004), available at http://www.dmso.mil/public/transition/vva; Pace, D. K., "The Referent in Simulation Development and Assessment," paper E04-SIW-024, in *Proc. 2004 European Simulation Interoperability Workshop (EuroSIW)*, available at http://www.sisostds.org.

[21]Balci, O., Nance, R. E., Arthur, J. D., and Ormsby, W. F., "Expanding Our Horizons in Verification, and Accreditation Research and Practice," in *Proc. 2002 Winter Simulation Conf.*, available at http://manta.cs.vt.edu/balci/papers/VVAHorizons.pdf.

[22]Pace, D. K., and Sheehan, J., "Subject Matter Expert (SME)/Peer Use in M&S V&V," an A4 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002).

[23]Metz, M., and Harmon, S., "Using Subject Matter Experts for Results Validation: A Progress Report," paper 02S-SIW-095, in *Proc. Spring 2002 Simulation Interoperability Workshop (SIW)*, available at http://www.sisostds.org/conference/View_Public_Number.cfm?Phase_ID-2.

[24]CJCSI 3170.01D, Joint Capabilities Integration and Development System (JCIDS) (12 Mar 2004), available at http://www.dtic.mil/cjcs_directives/cdata/unlimit/3170_01.pdf.

[25]Heitmeyer, C., "On the Need for *Practical* Formal Methods," invited paper, *5th Int. School and Symp.: Formal Techniques in Real Time and Fault Tolerant Systems (FTRTFT)*, pp. 18–26 (1998).

[26]Kuhn, D. R., Chandramouli, R., and Butler, R. W., "Cost Effective Use of Formal Methods in Verification and Validation," an A5 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002).

[27]Worley, D. R., Simpson, H. K., Moses, F. L., Aylward, M., Bailey, M., and Fish, D., *Utility of Modeling and Simulation in the Department of Defense Initial Data Collection*, paper D-1825, Institute of Defense Analyses (IDA) (May 1996), available at http://www.msiac.dmso.mil/ia/generaldocs.asp.

[28]Costs are not reported for the majority of the 359 models and simulations in 22 DoD acquisition programs. See Hillegas, A., Backschies, J., Donley, M., Duncan, R. C., and Edgar, W., *The Use of Modeling and Simulation (M&S) Tools in Acquisition Program Offices: Results of a Survey*, Hicks & Associates (31 Jan 2001).

[29]Kilikauskas, K. L., Brade, D., Gravitz, R. M., Hall, D. H., Hoppus, M. L., et al., "Estimating V&V Resource Requirements and Schedule Impact," a B4 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002).

[30]Examples of efforts to address uncertainties with more rigor and to facilitate scientifically defensible predictions from models and simulations include Ghanem, R., and Wojtkiewicz, S., *Uncertainty Quantification Short Course at the Soc. for Industrial and Applied Mathematics (SIAM) Conf. on Computational Science and Engineering* (9 Feb 2003), available at http://www.siam.org/meetings/cse03/UQShortcourse.htm; Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V., and Alvin, K. F., *Estimation of Total Uncertainty in Modeling and*

Simulation, Report SAND2000-0824, Sandia National Laboratories (Apr 2000); and *4th Int. Conf. on Sensitivity Analysis of Model Output (SAMO) 2004* (8–11 Mar 2004), available at http://www.floodrisknet.org.uk/events/SAMO%202004.

[31]Oberkampf, W. L., Trucano, T. G., and Hersch, C., "Verification, Validation, and Predictive Capability in Computational Engineering and Physics," a B1 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002). Subsequently published as Sandia Report SAND2003-3769 (Feb 2003).

[32]Menzies, T., "Verification and Validation and Artificial Intelligence," a B5 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002).

[33]Bigelow, J. H., and Davis, P. K., *Implications for Model Validation of Multiresolution, Multiperspective Modeling (MRMPM) and Exploratory Analysis*, Report MR-1750-AF, RAND Corp. (2003).

[34]For example, Lee, J. K., Lee, M. M., and Chi, S-D., "DEVS/HLA-Based Modeling and Simulation for Intelligent Transportation Systems," *SIMUL-T Soc. Mod. Sim.* **79**(8), 423–439 (1 Aug 2003); or Bunus, P., and Fritzson, P., "DEVS-Based Multi-Formalism Modeling and Simulation in Modelica," in *Proc. Summer Computer Simulation Conf.*, Vancouver, B.C., Canada, CD-ROM (2000).

[35]Giordano, J. C., abstract for a colloquium at the University of Virginia, *A Master's Project Presentation: Exploring the Constraints of Human Behavior Representation* (28 Apr 2004), available at http://www.cs.virginia.edu/colloquia/event391.html.

[36]Hughes, T., and Rolek, E., "Fidelity and Validity: Issues of Human Behavioral Representation Requirements Development," in *Proc. Winter Simulation Conf.*, New Orleans, LA, pp. 976–982 (2003).

[37]Harmon, S. Y., Hoffman, C. W. D., Gonzalez, A. J., Knauf, R., and Barr, V. B., "Validation of Human Behavior Representations," a B3 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002).

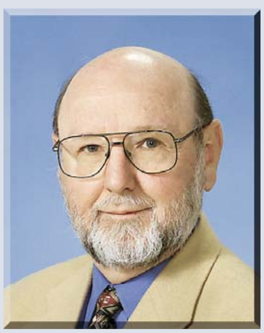
[38]MacKenzie, G. R., Schulmeyer, G. G., and Yilmaz, L., "Verification Technology Potential with Different Modeling and Simulation Development and Implementation Paradigms," an A1 paper in *Proc. Workshop on Foundations for Modeling and Simulation (M&S) Verification and Validation (V&V) in the 21st Century*, SCS, San Diego, CA, CD-ROM (2002).

[39]Theodorakis, M., Analyti, M. A., Constantopoulos, P., and Spyratos, N., "Contextualization as an Abstraction Mechanism for Conceptual Modeling," in *Proc. 18th Int. Conf. on Conceptual Modeling (ER99)*, available at http://www.ics.forth.gr/isl/publications/paperlink/ER99_paper.pdf.

[40]Others besides this author have drawn similar conclusions, though they have often been stated less bluntly. For example, *Modeling and Simulation in Defense Acquisition: Pathways to Success*, National Research Council (NRC) Board on Manufacturing and Engineering Design Committee on Modeling and Simulation Enhancements for 21st Century Manufacturing and Acquisition, available at http://www.dtic.mil/ndia/2002sba/castro.pdf.

[41]Available at http://www.cs.virginia.edu/~MaSTRi/. While the Modeling and Simulation Technology Research initiative (MaSTRi) vision shows appreciation for the reality that these areas are interrelated and MaSTRi plans to collaborate with the Illinois Institute of Technology Research Institute (IITRI) Modeling and Simulation Technology Research Institute, substantive endeavors have yet to be accomplished.

THE AUTHOR



DALE K. PACE studied at the University of Chicago, Capital Bible Seminary, and Luther Rice Seminary. He came to APL in 1963, leaving to be a correctional chaplain in the 1970s and returning in 1979. He was APL's liaison with the Naval War College (1987–1989) where, as an adjunct professor, he developed and taught an elective course on technology and naval warfare. At APL, Mr. Pace employs operations research (OR) methods in many areas of warfare analysis and contributes to OR methodology advances in scenario development, seminar war gaming, and modeling and simulation verification, validation, and accreditation (VV&A). He was Co-chair of the Military Operations Research Society (MORS) 1999 Simulation Validation Workshop (SIMVAL99) and Program Co-chair for the Foundations '02 V&V Workshop, and has been a member of the DoD VV&A Technical Working Group since its inception. His e-mail address is dale.pace@jhuapl.edu.