Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

PANEL: THE FUTURE OF RESEARCH IN MODELING & SIMULATION

Levent Yilmaz

Computer Science and Software Engineering
Auburn University
Auburn, AL 36849, USA

Simon J. E. Taylor

Modeling & Simulation Group
Department of Computer Science
Brunel University
Uxbridge. Middx, UB8 3PH, UK

Richard Fujimoto

School of Computational Science and Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

Frederica Darema

Mathematics, Information, Life Sciences Division
Air Force Office of Scientific Research
Arlington, VA 22203, USA

ABSTRACT

Due to the increasing availability of data and wider use of analytics, the ingredients for increased reliance on modeling and simulation are now present. Tremendous progress has been made in the field of modeling and simulation over the last six decades. Software and methodologies have advanced greatly. In the area of weather, future-casts based on model predictions have become highly accurate and heavily relied upon. This is happening in other domains, as well. In a similar vein, drivers may come to rely upon future-casts of traffic that are based on predictions from models fed by sensor data. The need for and the capabilities of simulation have never been greater. This panel will examine the future of research in modeling and simulation by (1) examining prior progress, (2) pointing out current weaknesses and limitations, (3) highlighting directions for future research, and (4) discussing support for research including funding opportunities.

1 INTRODUCTION

Today's application models and algorithmic methods employed in these models enable to represent, analyze, and predict behaviors of complex systems with unprecedented accuracy, and new modeling and simulation approaches can enable decision support systems thus advancing to role of modeling and simulation beyond the traditional analysis, and design or understanding roles of Modeling & Simulation (M&S), to having essential role in the operational cycle real time decision support on complex systems be they natural or engineered. In that context the presentations of the four panelists address the future of research opportunities and new capabilities from the following perspectives: Sustainable M&S Research for Science and Engineering (Yilmaz); Simulating Large Systems (Taylor); Parallel and Distributed Simulation (Fujimoto); and Dynamic Data Driven Applications Systems (Darema).

2 SUSTAINABLE M&S RESARCH FOR SCIENCE & ENGINEERING (LEVENT YILMAZ)

Solutions to scientific and engineering problems increasingly depend on the credibility and quality of simulation research.

978-1-4799-7486-3/14/$31.00 ©2014 IEEE

As the complexity of problems continues to grow, simulation-based science and engineering is becoming central to understanding, exploring, predicting, and influencing the behavior of artificial, physical, and natural systems. In a complex, dynamic, and evolving science-based innovation and knowledge landscape, the prominence and sustained vitality of the M&S discipline will rely on three critical pillars: (1) credibility, (2) relevance, and (3) accessibility of research.

2.1 The Credibility of M&S Research

According to the Merriam-Webster dictionary, *credibility* is the quality of being believed or accepted as true, real, or honest. Prior extensive research in verification and validation resulted in sound principles and guidelines for assuring the veracity of simulation results. However, growing credibility gap due to wide spread relax attitudes in communication of research artifacts is giving rise to the need for systematic reproducibility of simulation experiments and replication of models. Furthermore, disputes such as Climate Gate (Economist 2010) and article retractions due to unverified code and data (Chang et al. 2006; Alberts 2010) suggest a pressing need for greater transparency in M&S research.

As a fundamental principle of science, *reproducibility* aims to bring credibility and to instill confidence in research (Fomel and Claerbout 2009; Stodden 2010). It refers to the ability to recreate simulation results from existing simulation code and data. Specifically, in collaborative research involving remote teams reproducibility becomes a prerequisite. On the other hand, although exposure to simulation software is important, if done too early, it leads to the adoption of the assumptions of the original developer, resulting in the loss of independence necessary to replicate a model (Yilmaz 2013).

Scientific workflow systems (Anand et al. 2009; Oinn et al. 2004) and provenance-based tracking of research artifacts help improve reproducibility of experiments, and hence the adoption of these general purpose systems are expected to grow. In this context, the M&S community can play a critical role in developing proper standards for scholarly communication of reproducible simulation-based research. Specifically, independent replication of models, which is a weakness of existing workflow systems, can be addressed by recent developments in Model-Driven Engineering, which facilitates automated transformation of models across formalisms, to support cross-validation, reuse, and model longevity. Besides, stochastic simulation experiments involve strategic (e.g., design of experiments) and tactical plans (e.g., how many runs needed) that are often implicit and hence cannot be reliably reproduced, unless explicitly specified and managed. Therefore, to support their management, tools and environments are necessary to assist the overall simulation experiment lifecycle, which includes design of experiments, automated synthesis and deployment of experiment scripts, aggregation and analysis of data, and refinement and online adaptation of experiment designs through feedback as learning takes place.

The provision of such automated tools will help improve the state of the art and practice in replicability of models and reproducibility of simulation experiments. In the short term, however, authors can provide hyperlinks to simulation code and data, or use open-source environments such as Source-Forge or Github for transparency. To facilitate publishing, citing, maintaining, and discovering research data, authors can use infrastructures such as the DataVerse Network Project (http://thedata.org). Funding opportunities can incentivize further development and adoption of research practices that streamline reproducibility. For instance, funding agencies can support research groups that implement reproducible research to facilitate discerning the information requirements and tools that improve the practice of model replication and simulation experiment management. Also, funding that support formation of research communities and communities of practice can sustain maintainable and reproducible research.

2.2 The Relevance and Accessibility of M&S Research

Research is a creative endeavor that requires both novelty and usefulness. As discussed in (Yilmaz and Smith 2008), the prominence of M&S in scientific and engineering research requires proper alignment of the discipline, the field, and the stakeholders, including the practitioners. The academic M&S literature is extensive and full of rigor, which is important for diligence and precision. The rigor and diligence brought by the academic language often requires practitioners to wade through pages of information to access and discern relevant information. However, communicating research results to stakeholders in an accessible language using the right platform will be mutually beneficial to both the practitioners and the researchers. Contextualizing solutions and ensuring that research is relevant and reflects the present and future needs will require publishing evidence-based short papers using non-academic language while validating the findings in a specified context.

Practitioners often consult with books, blogs, forums, online video tutorials, forums, and brief one-to-two page experience reports. To bridge the theory and practice divide, new communication channels and collaboratories are needed to improve the communication gap. For instance, the monthly Newsletter of the Society for Modeling and Simulation International was recently rejuvenated to provide researchers and practitioners a medium to discuss research problems and technical solutions in a practical context. Researchers may venture into various alternative dissemination channels such as wikis, social networks, and blogs to improve the practical utility and accessibility of their research. Research publications are critical to demonstrate the reliability, validity, and integrity of results; however, decision-makers and practitioners often need best practices, prescriptive guidance, and customizable generic solutions that are effective in addressing well-defined problems.

M&S scholarship equipped with the tools of the digital age, is expected to improve the *compatibility*, *trialability*, *observability*, and *traceability* of research artifacts. Compatibility refers to portability or accessibility of simulation artifacts on technical platforms for use by others that aim to reproduce the results of a study. Trialability ensures repeatability of experiments and generating the data. Observability is needed to not only introspectively access artifacts to improve understandability, but also to examine the outputs for comparison to reference data published in the manuscript. Transparency reflects the responsibility of scholars to register simulation artifacts along with the publication to enable legitimization and access by others. Among the incentives for transparency are reputation, visibility, and developing a community of practice around the artifacts produced during research. Journals and funding agencies can play a significant role by implementing policies that encourage provision of models and code, while requiring reproducibility reviews prior to publication. Maintenance of reproducible simulation artifacts for the purpose of preserving and persistent availability is essential to facilitate their continuous use. As the M&S community, we need to pay attention to not only the technical infrastructure, but also social and legal context. Social context focuses on the formation of community of practice around artifacts and governance of their further development and maintenance. Legal context involves intellectual property management schemes and access rights. As we extend the relevance and access of our research products, it will be necessary to institute copyright management mechanisms for data and simulation code, as well as the media components and the published manuscript. The use of open licensing schemes adopted by the Open-Access and Open-Data projects under the Creative Commons (http://creativecommons.org) framework can secure intellectual property rights, while maximizing access and citation to various types of digital research artifacts.

3 SIMULATING LARGE SYSTEMS (SIMON TAYLOR)

A "large" system may be one in which aspects of the scope, complexity, size and of the system in terms of structure and/or data make it challenging to model and to simulate. For example, a coronary system might involve a range of different subsystems representing different aspects of the heart from fluid flow, to molecular processes and disease progression. A healthcare pathway system may consist of many different subsystems representing different stages of a patient's primary and secondary care. A supply chain may consist of many different manufacturing systems and economic systems that represent the supply and demand of a joint enterprise. Modeling a large system might use hybrid techniques that combine discrete and continuous methods to create submodels representing different subsystems or functions that are combined to make the model. Some of these submodels might already exist and could be reused. Models would be developed using different M&S software and languages. Large system models might require access to many potentially large datasets. Simulating a large model could be very demanding in terms of computing power. A single run of the simulation might run very slowly. Even if a single run ran in an acceptable time, many runs would be needed during verification and validation as well as during experimentation. Also, it would be surprising if a large model was created by a single person. It is reasonable to assume that large models would be developed by a team of modelers (potentially many teams of modelers). How could these modeling teams and their development efforts be coordinated?

In many ways the above mirrors experiences of researchers in some scientific areas such as high energy physics, astronomy and bioinformatics. In these researchers work together in international multidisciplinary teams (sometimes called virtual research communities (VRCs) or communities of practice (COPs)) towards their goals. These can involve the creation of large multi-scale, multi-paradigm models which demand huge amounts of computing power to simulate them. Their work is supported by an integrated framework of Information and Communication Technologies (ICT) termed *e-Infrastructures* or *Cyberinfrastructures*. A typical e-Infrastructure architecture is shown in figure 1. From the bottom up high performance networks such as GEANT (www.geant.net), along with other commonly accessible networks (e.g. the Internet) support the high speed transfer of data between e-Infrastructure facilities. These include Distributed Computing Infrastructure facilities (such as Grid Computing, Cloud Computing or dedicated specialist High Performance Computing platforms), data infrastructures for storage and curation, and sensor and instrumentation networks. Almost complete single sign-on access to these are provided by an Authentication and Authorization Infrastructure supported by Certification Authorities and Identity Federations.

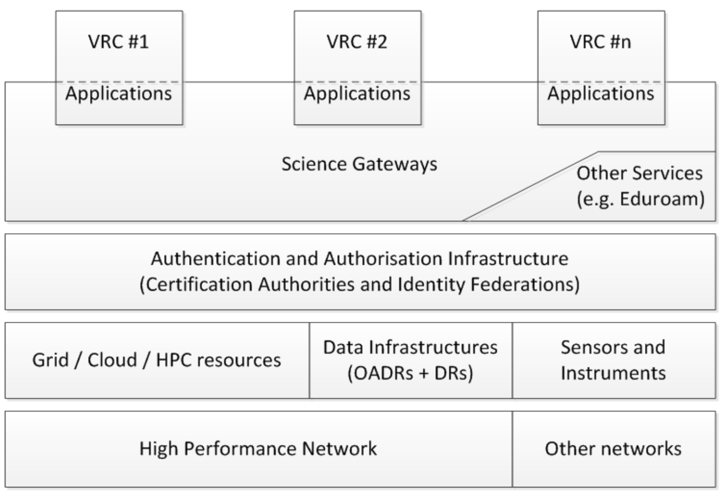

Figure 1: A Typical e-Infrastructure Architecture (*eI4Africa.eu*).

VRC access is supported by a science gateway that provides convenient access to applications and services deployed across the potentially worldwide facilities of an e-Infrastructure. A strong element of this is that all the e-Infrastructure services are supported by a range of standards and services (see https://wiki.egi.eu/wiki/Standards for an overview of standards in the area; Eduroam is an example of a worldwide service that provides roaming network access across research and education networks). In Europe e-Infrastructures are developed and maintained by EGI.EU with support from the European Commission and member nations. Other initiatives are promoting and developing e-Infrastructures in other regions and other sectors. The eI4Africa project (eI4Africa.eu), for example, is developing science gateways for African-based VRCs (see the African Science Gateway https://sgw.africa-grid.org/). The

CloudSME project uses a similar architecture with elements of workflow specification to develop e-Infrastructures for simulation end users in manufacturing and engineering (www.cloudsme.org).

Large international networks of scientists are benefiting from advanced ICT infrastructures that are enabling them to make scientific breakthroughs that would arguably be impossible without this technological support. It could be argued that our M&S community is made up of many VRCs that contribute to the advancement of science and industry. As outlined above, the needs of modeling and simulating large systems has a parallel to the experiences of international communities of scientists. As the last of these projects indicate, some work is in progress to link these two "worlds" together. However, there is much more work to be done if our M&S communities are ever to take on the challenges and to create the clear benefits of modeling and simulating larger and more complex models.

4 PARALLEL AND DISTRIBUTED SIMULATION (RICHARD FUJIMOTO)

Parallel and distributed simulation is concerned with the technologies associated with distributing the execution of a single run of a discrete event simulation program across multiple processors (Fujimoto 2000). Parallel discrete event simulation (PDES) focuses on high performance computing systems while distributed simulation is concerned with exploiting distributed computing platforms that can cover a much broader geographic extent ranging from machines interconnected through a local area network to globally distributed computers communicating via the Internet. While the central goal of PDES is to accelerate the execution of the simulation, the goals of distributed simulations are often broader, and can include objectives such as reuse of existing simulations or exploitation of geographically distributed resources such as equipment or people that are difficult or costly to co-locate.

Early work in PDES focused on the synchronization algorithm that is required to ensure that the parallel execution of the simulation produces exactly the same results as a sequential execution on a single processor. The parallel and distributed simulation field began in the late 1970's with seminal work by Chandy, Misra, and Bryant who defined the synchronization problem and a solution approach that has come to be known as conservative synchronization (Bryant 1977, Chandy and Misra 1979). An alternative approach known as optimistic synchronization was developed in the 1980's, originating with seminar work by Jefferson who developed the Time Warp algorithm (Jefferson 1985). The study of PDES synchronization algorithms flourished in the 1980's and 1990's.

At the same time, much of the work in distributed simulation originated with the SIMNET project in the 1980's that focused in interoperability among simulation (Miller and Thorpe 1995). The field developed through subsequent efforts both in technology development and standardization, included efforts such as Distributed Interactive Simulation (IEEE Std 1278.1-1995 1995, IEEE Std 1278.2-1995 1995) the Aggregate Level Simulation Protocol (Wilson and Weatherly 1994), and the High Level Architecture standard (IEEE Std 1516-2010 2010).

4.1 Parallel and Distributed Simulation Today

The field has enjoyed many impressive technical successes over the years. Numerous case studies have demonstrated the ability of PDES technology to accelerate the execution of discrete event simulations. For example, (Fujimoto et al. 2003) examined packet-level simulation of computer communication networks on supercomputers. Experiments yielded performance as high as over 200 million events processed per second using a conservative synchronization algorithm executing on a supercomputer using 1,536 processors. By comparison, comparable simulators executing on a sequential machine yielded performance less than 200,000 events per second. Later studies using a synthetic benchmark called PHOLD yielded performance exceeding 529 million events per second using an optimistic synchronization algorithm on a 16,384 processor IBM Blue Gene machine (Perumalla 2007). Recently, Barnes et. al were able to achieve 504 billion events per second using almost 2 million cores of a Blue Gene/Q machine (Barnes et al. 2013).

It may be noted that these three studies yielded event rate performance *per core* of 138K events/second/core (in 2003), 32K (in 2007), and 256K (in 2013), representing only a factor of 2 improvement in single core performance over the last 10 years. These data highlight the fact that performance improvements are being driven almost entirely by increases in parallelism. This is not surprising. Processor clock rates have seen only modest increases since 2005 due to physical constraints concerning heat dissipation, resulting in an explosion in the number of cores in supercomputer architectures since 2005. Throughout much of the 1990's and up until 2005 the most powerful supercomputers contained only thousands of cores. The most powerful machines today contain millions.

Despite these successes, PDES technology has yet to penetrate the broader modeling and simulation industry. The vast majority of discrete event simulation executed today, even those modeling large-scale systems, execute on sequential computers. In (Fujimoto 1993) several challenges that stood in the way of widespread adoption of PDES technology were presented. Among the avenues that were proposed to achieve wider acceptance, the application library and automated parallelization approaches have perhaps seen the most progress. The application library approach is exemplified by Qualnet (Scalable Network Technologies 2012) that enables one to configure a network simulation from a graphical user interface while hiding the details of parallel execution in an underlying simulation executive. In the automated parallelization arena, the "self-federating" approach where multiple instantiations of a single sequential simulation are created and interconnected, with each modeling a portion of the overall system has emerged as a practical approach to realizing parallel simulations. For example, each sequential simulation might model a sub-network of a large telecommunications system. Examples of this approach include (Nicol and Heidelberger 1996, Riley et al. 2004) among others. Because this approach effectively involves federating a sequential simulation with itself in a parallel or distributed computing environment, it avoids many of the interoperability issues that arise when federating different simulators.

Similarly, the penetration of distributed simulation technology into the broader modeling and simulation industry still has a long ways to go. (Strassburger, Schulze et al. 2008) reports the findings of a peer study in the context of distributed virtual environments that indicate the technology is widely believed to have the potential for broad impact, but widespread adoption in industry has not yet materialized. Massive, multiplayer on-line game systems represent one area where the technology has seen extensive commercial use. Nevertheless, widespread adoption of the technology does not appear to have reached its fullest potential. In (Taylor, Mustafee et al. 2009) the authors articulate the need for new standards to help address this issue by increasing interoperability of simulation models and tools.

4.2 Challenges

Below we describe several challenges to help guide the parallel and distributed simulation field in directions that we believe will significantly increase the impact of this technology.

4.2.1 Scalable PDES Simulations of Realistic Networks

A substantial amount of effort in the past decade has focused on the topology of networks that arise in real-world applications. One class of topologies that have become known as scale-free networks (Barabasi and Albert 1999) have been observed to arise in many different applications. A scale-free network is one where the node degree follows a power law distribution. A distinguishing characteristic of scale-free networks is a significant number of nodes, referred to as *hub* nodes, contain a large node degree, while most nodes, often referred to as *leaf* nodes, contain relatively small degree. This is in stark contrast to the regular, symmetric networks typically used in large-scale PDES performance studies thus far. Such networks often have high degree hub nodes that turn into bottlenecks for PDES computations. Scale-free networks have received a considerable amount of attention in recent years because it has been observed that many real-world systems contain networks that exhibit the scale-free property (Wang and Chen 2003). For example, it is widely believed that the autonomous system (AS) level topology of the Internet is scale-free (Faloutsos et al. 1999, Faloutsos et al. 2003, Zhang et al. 2011). In the broad area of systems biology, the study of complex biological systems, protein-protein interaction networks have been demonstrated to follow scale-free distributions (Kitano 2002). Some financial networks such as the interbank payment network exhibit scale-free behavior (Soramaki et al. 2007). Social networks, the world-wide-web, the internal structuring of superconductors, the airline transportation network, and human interaction networks such as models of the spread of diseases have been reported to exhibit scale-free properties.

The network topology has a large impact on the efficiency of parallel discrete event simulation techniques in terms of parallelism and overhead of the synchronization protocol. It has been observed empirically that the distribution of event-level parallelism in simulations of communication networks can lead to severe load imbalances (Liu and Chien 2004). In (D'Angelo et al.) load distribution issues in scale-free network simulations are examined. The relationship between power law topology and parallel simulator performance was studied in (Pienta and Fujimoto 2013) where it was observed through both analytical models and simulations that very large network simulations may yield very limited parallelism. For example, scale-free networks containing tens of thousands of network nodes often exhibit less than hundred-fold parallelism, suggesting limited opportunity for existing approaches to PDES to accelerate performance in many real-world applications. The limited amount of parallelism in these networks coupled with the reliance on massive parallelism to achieve high performance on modern supercomputers makes exploitation of PDES challenging for many real-world applications.

4.2.2 PDES Benchmark Programs

One modest, albeit important challenge lies in defining new, realistic benchmark applications to evaluate new PDES technologies. Since its early years the PDES community has been utilizing a benchmark program called PHOLD as a means to evaluate performance (Fujimoto 1990). While useful in the early years of the field, and of some practical use today, PHOLD fails to capture many critically important aspects of real PDES applications. PHOLD owes its origins to another benchmark called the HOLD model that was used to benchmark priority queue implementations in sequential discrete event simulations (Jones 1986). Most uses of PHOLD use regular topologies such as a toroid or a fully connected network, with the message sent to a neighboring LP with any neighbor equally likely to be selected. PHOLD has the virtue that it is very easy to implement, which likely accounts for its popularity. However, PHOLD, has a number of important deficiencies that can lead to very misleading performance results. The first, obvious limitation is a toroid or fully connected network topology leads to highly symmetric and regular networks with well balanced computation workloads. The benchmarking studies that were described earlier that reported impressive performance results on supercomputers all utilized highly regular networks. Such topologies are very different from real-world networks that are typically irregular, with skewed degree distributions, or exhibiting scale-free properties. Further, PHOLD does not incorporate events with different computational requirements or different dynamics such as varying numbers of events scheduled from one event to the next. We believe new benchmarks that retain PHOLD’s easy to implement characteristic while capturing more realistic applications are needed to help further develop PDES technology.

4.2.3 Large-Scale PDES Solving Grand Challenge Problems

Large-scale PDES performance studies to date have largely been technology demonstrations rather than use of PDES to solve specific real-world problems. While technology demonstrations are important and serve a purpose, the real, lasting impact of the technology will come from its use to solve real-world problems that cannot otherwise be solved. With only a few exceptions, the vast majority of commercial simulations are sequential. For PDES to have a substantial impact in the broader modeling and simulation community, several technical challenges must be overcome. PDES must be largely transparent to the simulation user. This requires simple simulation languages that are natural to program and readily expose parallelism to the underlying simulation engine. The self-federating approach described earlier offers at least a partial solution to this problem. Important issues such as model partitioning, load distribution, and synchronization must be handled automatically, and effective debugging environments must be readily available. Output analysis must be no more difficult than that corresponding to sequential execution.

4.2.4 Practical, Large-Scale Simulation in the Cloud

Cloud computing services such as Amazon Web Services' Elastic Compute Cloud (EC2), Microsoft's Azure platform, and Google's AppEngine provide virtualized hardware and software that can be accessed via the web. It potentially addresses a significant impediment to the widespread adoption of PDES technology by lowering the barrier to gaining access to high performance computing machines. The cloud's "pay-as-you-go" economic model eliminates the need to purchase, operate and maintain high performance computing equipment locally (Fujimoto et al. 2010). Further, by providing parallel and distributed simulation software as a service, cloud computing offers the ability to hide many of the complications of executing parallel and distributed simulation codes from the user, offering the potential to make exploitation of this technology much less risky than is the case today.

However, exploitation of PDES on the cloud introduces new challenges. Preliminary work in benchmarking parallel scientific programs in Amazon's EC2 observed that parallel scientific codes executed over EC2 ran significantly slower compared to execution on dedicated nodes of a cluster (Walker 2008, Ekanayake and Fox 2009). Two issues are communications and interference. Cloud environments are often better at providing high bandwidth communications among applications than in providing low latency, and high delay variance has often been observed in practice (Walker 2008). This is problematic for many simulation applications that are accustomed to sending many small messages requiring quick delivery rather than fewer large messages requiring high bandwidth alone. Further, cloud environments are shared among many users, and individual users are not guaranteed exclusive access to the processors assigned to that user's virtual cluster. In principle, gang scheduling techniques can be used to ensure an individual user is allocated a set of physical nodes at the same time instant, however, this property may not be guaranteed by the cloud provider. This can lead to difficulties for parallel simulation applications, especially those that utilize optimistic synchronization techniques.

4.2.4.1 Real-Time Dynamic Data Driven Parallel and Distributed Simulation

The ability to effectively exploit cloud computing for PDES applications greatly facilitates exploitation of the technology in real-time (or on-line) simulation applications. By this we mean the use of simulation to optimize operational systems while they are running. Also known as symbiotic simulations (Fujimoto et al. 2002) or dynamic data driven application systems (DDDAS) (Darema 2004), these systems use networks of fixed and/or mobile sensors, possibly coupled with the use of crowd-sourced data to develop a model of the current state of the system. Simulations then utilize this captured state to project future states and perform what-if analyses, often in conjunction with other optimization tools and techniques to inform operators as to how to improve system performance, or reconfigure the sensor network to improve system performance. Numerous applications of this technology exist, such as optimizing supply chains, transportation systems, communication networks and energy systems, among others.

5 DYNAMIC DATA-DRIVEN APPLICATION SYSTEMS (FREDERICA DAREMA)

This section discusses new directions in modeling and simulation, namely the Dynamic Data Driven Applications Systems (DDDAS) paradigm (Darema 1990; Darema 2000).

5.1 Definition of DDDAS

DDDAS is a paradigm whereby selected on-line instrumentation-data (or archival data) are dynamically integrated into an executing model/simulation in a feed-back control loop, with the executing model in reverse guiding adaptively the instrumentation processes. The effect of this paradigm is to create more efficient and more accurate modeling and simulation capabilities by either using the actual data to compensate for aspects of the system not captured accurately in the model/simulation, or by using the actual data to replace targeted parts of the computation in order to speed the modeling/simulation process and which can also result of enabling real-time decision support capabilities with the accuracy of full scale simulation; and in reverse create capabilities for more efficient and effective instrumentation capabilities through the modeling/simulation driven control of the instrumentation processes, for example, collect data in targeted and adaptive ways. Inherently, the DDDAS feed-back control loop unifies complex computational modeling of a system with the real-time data acquisition and control aspects of the system. DDDAS is also referred to as InfoSymbiotics or InfoSymbtiotic Systems (in the mid-eighties to early 2000 was also referred to as Gedanken Laboratory).

5.2 Drivers and Changing Landscape in M&S

Application systems today and those foreseen in the future, be they natural, engineered, or societal, have unprecedented scales of complexity, interconnectivity, and interdependence, across components of a system as well as across systems. Such complex systems require more advanced methods for analysis, understanding, design, and management. The methods needed go beyond the static modeling and simulation methods of the past. New approaches such as DDDAS augment and enhance system models/simulations through continually updating critical regions of the solution space of the problem with information from monitoring and control/feedback aspects of the system. The needs for autonomic capabilities and optimized management of engineered systems, consisting of heterogeneous and dynamic components and resources makes more compelling the need for new methods such as DDDAS, not only for the design stage of engineered systems but also for managing the operational cycle of such systems.

The emergence of several technological and methodological advances over the last ten years has resulted into an added impetus for exploiting the integration of modeling with observation and actuation as envisioned in the DDDAS paradigm, making DDDAS more timely than ever. Such advances include the increasing emphasis in multi-scale and multi-modal modeling – in DDDAS multiple scales may be invoked dynamically based on the dynamic data inputs. In addition, the emerging multicore-based computing technologies are transforming the computational capabilities in the high-end computing as well as the real-time data acquisition and control systems, with the concomitant emphasis ubiquitous sensing and control instrumentation capabilities. Furthermore there is also tremendous increase in networking capabilities for streaming large volumes of data remotely and connecting multiple distributed and heterogeneous data and computation resources. In tandem with that, there is an emerging emphasis of comprehensive cyberinfrastructures to support complex systems.

There is a triad of broad approaches used to understand and analyze the behavior of systems, be they natural or engineered, namely 1) theory – the theoretical concept &/or principles about a system; 2) the mathematical representation of these theoretical principles about the system; and 3) the experimental probing of the system through instrumentation. Traditional approaches in modeling and simulation consider this triad of theory, M&S, and instrumentation as distinct and serially related; that is: we create a theoretical concept about a system, then we represent this theoretical concept through a model or sets of models (here model is used to denote either model or simulation), and we use the instrumentation data as inputs to the model and the computation results present mathematically a state (or set of states) of the system; subsequently we may launch additional computations of the model with other sets of data inputs, to compute other states of the system; based on these results, we may also change the theoretical concepts about the system and the corresponding models. DDDAS changes the static and serialized relation between modeling/simulation of a system and the instrumentation of the system, to a dynamically integrated one.

5.3 New Directions through DDDAS:

With DDDAS, in addition to the initial data inputs, also termed as "static data inputs", as an application model executes we incorporate dynamically additional data into the model – these data may be data acquired on-line in real-time or they can be data that have been previously acquired (archival data), and selected such data, are incorporated into the executing model, in selected parts of the phase space of the problem, as dictated by the executing model. Examples include replacing parts of the model computation with actual data to speed-up the modeling process, or use selected data to complement the model for the aspects of the system not captured at all or not captured accurately by the model. Data assimilation is for example a special case of DDDAS, where actual data are used to constrain the size of the error bars (data uncertainties) in the solution vector of the PDE (Partial Differential Equation) representing the system; as the computed data in the solution vector and their corresponding computed errors propagate through the simulation (the PDE solver), computed data with large uncertainties are replaced with actual data that have smaller error bars, and computation restarts with the resulting updated vector. In addition in DDDAS the model- driven instrumentation-control entails that measurement data can be selectively targeted to collect the data useful to improve the accuracy of the model or to speed-up the model execution. Thus the DDDAS methods result into more efficient data measurement capabilities. Moreover in the case where measurement data result from multiple instrumentation resources (such required heterogeneous sensor networks) the traditional approaches select such data in static and ad-hoc ways; DDDAS allows dynamic and adaptive scheduling of the data collection and dynamic management and control of such instrumentation data resources. Thus the DDDAS-based methods allow more efficient and effective utilization of such heterogeneous sensor and controller resources.

5.4 DDDAS and Big Data - Big Computing:

In DDDAS the computational and instrumentation aspects of an application system become a unified representation of the system, and the underlying computational and instrumentation platforms become a unified platform which encompasses and may span the range of platforms, from high-end and mid-range computing to the real-time data acquisition and control, and to hand-held personal devices. In DDDAS and with the advent of ubiquitous sensing and control, Big Computing spans beyond the exascale to include the computing on the multitudes of heterogeneous sensors and controllers; so here, when we refer to Big Computing will refer to the computing at the high-end plus the highly distributed computing on the collection of multitudes sensors and controllers, all-together acting as a unified platform. Likewise, in DDDAS the notion of Big Data spans beyond the data generated by large scale computations and large instruments to also include the collection of data from the multitudes of sensors and controllers. In an era where we speak about Big Data and the "data deluge" challenge, the DDDAS paradigm where an executing model dynamically and adaptively manages the instrumentation of a collection of heterogeneous sensors and controllers, creates new capabilities that go beyond the traditional static and ad-hoc ways of managing such resources, and allows efficient management of such resources. DDDAS enables to collect data targeted and selective ways, as dictated by the related application system needs; in other words DDDAS mitigates the "data deluge" by allowing to collect and manage data in "smart ways".

5.5 DDDAS Technical Challenges and Opportunities for New Capabilities

DDDAS is a compelling paradigm, and efforts enabling the capabilities sought under the rubric of DDDAS span several dimensions, requiring synergistic multidisciplinary research. DDDAS drives innovations in application modeling and simulation methods where executing models/simulations can incorporate dynamically additional data, where other application models can be dynamically invoked based on the dynamic data inputs (for example multiple scales of models and models of multiple modalities of the system), and where in return interface with the instrumentation systems and control the measurement and actuation processes in these systems. In addition the models and the algorithms used in these models need to be tolerant in their stability and have guaranteed convergence properties when subject to perturbations from the dynamically incorporated data into the models and the associated algorithms, and need new algorithmic methods for efficient uncertainty quantification and efficient estimation of error propagation across dynamically invoked application models. Other advances needed include new system software methods to support the dynamic and adaptive runtime requirements of such applications which not only entail heterogeneous underlying resource support (spanning for example from high-end computing to real-time), but also where the computational, memory, communication, and I/O requirements of such applications change depending on the dynamic data inputs. In addition, DDDAS environments require new methods of interfaces to and management of instrumentation resources for example dynamic and adaptive optimized management of heterogeneous collections of networks of sensors and/or networked controllers, and intelligent methods of large scale heterogeneous data management.

The approaches discussed here enable new capabilities for more accurate modeling methods for analysis and understanding of natural and engineered systems, and in the design and operational management of engineered systems, decision support methods with the accuracy of full-scale simulation models, and more efficient instrumentation and control methods. All these create unprecedented opportunities for creating and exploiting complex engineered systems, understanding societal systems, and new ways of understanding natural systems and responding to natural events. Examples of advances and new capabilities include analysis and decision support for structural systems (Oden et al. 2012; Bazilevs et al. 2013; Allaire et al 2014), medical systems (Fuentes et al. 2013), environmental systems (Patra et al. 2014; Douglas et al. 2006; Patrikalakis et al. 2004), and critical infrastructure systems, such as transportation systems (Fujimoto R. M. 2004) and electrical powergrids (McCalley et al. 2007); Celik et al. 2010) as well as many other application areas (AFOSR DDDAS Program, September-October 2013 PI Meeting; NSF Workshops 2000, 2006; AFOSR/NSF Workshop 2010; DDDAS Community Webpage www.dddas.org).

REFERENCES

AFOSR DDDAS Program PI Meeting. 2013. https://community.apan.org/afosr/w/researchareas/10782.dddas-program-pi-meeting.aspx

AFOSR and NSF. 2010. *Jointly Sponsored Mutli-Agency DDDAS Workshop*. 2010. www.dddas.org

Alberts, B. 2010. "Editorial Expression of Concern." *Science*, 327, 144. http://www.sciencemag.org/content/327/5962/144.1.full.

Anand, M., S. Bowers, T. McPhillips, and R. Ludascher. 2009. "Exploring Scientific Workflow Provenance Using Hybrid Queries over Nested Data and Lineage Graphs." In *Proceedings of the Scientific and Statistical Database Management*, 237–254.

Allaire, Allaire, D., D. Kordonowy, M. Lecerf, L. Mainini, and K. Willcox. 2014. "Multifidelity DDDAS Methods with Application to a Self-aware Aerospace Vehicle." In *Proceedings of the 2013 International Conference on Computational Science*. Elsevier.

Barabasi, A. L. and R. Albert 1999. "Emergence of Scaling in Random Networks." *Science* 286(5439): 509-512.

Barnes, P. D., C. D. Carothers, D. R. Jefferson and J. M. LaPre. 2013. "Warp Speed: Executing Time Warp on 1,966,080 Cores. "*Principles of Advanced Discrete Simulation***.** 327-336.

Bazilevs, Y., M-C. Hsu, and M. T. Bement. 2013. "Adjoint-based Control of Fluid-Structure Interaction for Computational Steering Applications." In Proceedings of *the 2013 International Conference on Computational Science*. Elsevier.

Bononi, L., G. D'Angelo and L. Donatiello. 2003. “HLA-Based Adaptive Distributed Simulation of Wireless Mobile Systems.” In *Proceedings of the 17th Workshop on Parallel and Distributed Simulation*: 40-49.

Bryant, R. E. 1977. Simulation of Packet Communications Architecture Computer Systems. *MIT-LCS-TR-188*.

Celik, N. 2010. “DDDAS-based Multi-fidelity Simulation Framework for Supply Chain Systems*.” IIE Transactions* V.42(5):325-341.

Chandy, K. M. and J. Misra. 1979. "Distributed Simulation: A Case Study in Design and Verification of Distributed Programs*." IEEE Transactions on Software Engineering*. SE-5(5): 440-452.

Chang, G., B. C. Roth, L. C. Reyes, O. Pornillos, Y-J. Chen, P. A. Chen. 2006. “Retraction”, *Science*, 314, 1875, http://www.sciencemag.org/cgi/content/full/314/5807/1875b.

Darema, F. 1990. “Parallel Applications and the Gedanken Laboratory.” Presentation *at Conference of the Society of Engineering Sciences*, SanteFe, NM.

Darema, F., Douglas, C., Deshmukh, A (Eds.) 2000. *Report of the NSF sponsored workshop on Dynamic Data-Driven Applications Systems*, March 8-10,2000, Ballston, VA. http://www.nsf.gov/cise/cns/dddas/

Darema, F. 2004. “Dynamic Data Driven Applications Systems: A New Paradigm for Application Simulations and Measurements.” *International Conference on Computational Science.*

Darema, F. 2005. “Dynamic Data Driven Applications Systems (DDDAS): New Capabilities for Application Simulations and Measurements.” In *Proceedings of the ICCS05*.

Darema, F. 2009. “DDDAS Computational Model and Environments.” *Journal of Algorithms and Computational Technology*.

Douglas, C. C. et al 2006. “DDDAS Approaches to Wildland Fire Modeling and Contaminant Tracking.” In *Proceedings of the 2006 Winter Simulation Conference*.

Economist. 2010. “The Clouds of Unknowing.” *The Economist*, March 18, 2010, http://www.economist.com/node/15719298.

Ekanayake, J. and G. Fox. 2009. *High Performance Parallel Computing with Clouds and Cloud Technologies. Bloomington*, IN, Department of Computer Science, Indiana University.

Faloutsos, M., P. Faloutsos and C. Faloutsos. 1999. "On power-law relationships of the internet topology*." SIGCOMM Computer Communications Review*. 29(4): 251-262.

Fomel, S. and J. Claerbout. 2009. “Guest Editors’ Introduction: Reproducible Research.” *Computing in Science and Engineering*, 11, 5–7.

Fuentes, D., Oden, J.T.; Diller, K. R.; Yung, J.; Feng, Y.; Hazle, J. D.; Shetty, A.; and Stafford, R. J. Stafford. 2013. “Computational and MR-guided Patient-Specific Laser Induced Thermal Therapy of Cancer”, *ICES Report* 13-33, The Institute for Computational Engineering and Sciences, The University of Texas at Austin.

Fujimoto, R. M. 1990. “Performance of Time Warp Under Synthetic Workloads.” In *Proceedings of the SCS Multiconference on Distributed Simulation*.

Fujimoto, R. M. 1993. "Parallel Discrete Event Simulation: Will the Field Survive?" *ORSA Journal on Computing* 5(3): 213-230.

Fujimoto, R. M. 2000. *Parallel and Distributed Simulation Systems*, Wiley Interscience.

Fujimoto, R. M., D. Lunceford, E. Page and A. Uhrmacher (Eds.). 2002. *Grand Challenges in Modeling and Simulation*, Technical Report 350, Schloss Dagstuhl, Seminar No. 02351.

Fujimoto, R. M., et al. 2004. “Dynamic Data-driven Application Simulation of Surface Transportation Systems.” *ICCS 2004, Part III, LNCS* 3993, Springer-Verlag.

Fujimoto, R. M., A. W. Malik and A. J. Park. 2010. "Parallel and Distributed Simulation in the Cloud." *SCS Modeling and Simulation Magazine*, Society for Modeling and Simulation, Intl. 1(3).

Fujimoto, R. M., K. Perumalla, A. Park, H. Wu, M. H. Ammar and G. F. Riley. 2003. “Large-Scale Network Simulation: How Big? How Fast?” *Symposium on Modeling, Analysis, and Simulation of*

Computer and Telecommunication Systems.

IEEE Std 1278.1-1995. 1995. *IEEE Standard for Distributed Interactive Simulation -- Application Protocols*. New York, NY, Institute of Electrical and Electronics Engineers, Inc.

IEEE Std 1278.2-1995. 1995. *IEEE Standard for Distributed Interactive Simulation -- Communication Services and Profiles*. New York, NY, Institute of Electrical and Electronics Engineers Inc.

IEEE Std 1516-2010. 2010. *IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA) -- Framework and Rules*. New York, NY, Institute of Electrical and Electronics Engineers, Inc.

Jefferson, D. 1985. "Virtual Time." *ACM Transactions on Programming Languages and Systems* 7(3): 404-425.

Jones, D. W. 1986. "An Empirical Comparison of Priority-Queue and Event-Set Implementations." *Communications of the ACM* 29(4): 300-311.

Kitano, H. 2002. "Systems Biology: A Brief Overview." *Science* 295(5560): 1662-1664.

Liu, X. and A. A. Chien. 2004. "Realistic Large-Scale Online Network Simulation." In *Proceedings of the 2004 ACM/IEEE Conference on Supercomputing*.

McCalley, J., et al. 2007. "Integrated Decision Algorithms for Auto-steered Electric Transmission System Asset Management." In *Proceedings of the 2007 International Conference on Computational Science*. Springer.

Miller, D. C. and J. A. Thorpe, 1995. "SIMNET: The Advent of Simulator Networking." In *Proceedings of the IEEE* 83(8): 1114-1123.

Nicol, D. M. and P. Heidelberger. 1996. "Parallel Execution for Sequential Simulators." *ACM Transactions on Modeling and Computer Simulation* 6(3): 210-242.

NSF DDDAS Workshops 2000 and 2006. http://www.nsf.gov/cise/cns/dddas/index_wkshp.jsp and http://www.nsf.gov/cise/cns/dddas/2006_Workshop/index.jsp.

Oden, J.T., et al. 2012. "Dynamic Data Driven Application Systems for Monitoring Damage in Composite Materials Under Dynamic Loads." *ICES Report* 12-37, The Institute for Computational Engineering and Sciences, The University of Texas at Austin.

Oinn, T., M. Addis, J. Ferris, D. Marvin, M. Senger, M. Greenwood, T. Carver, and K. Glover. 2004. "Taverna: A Tool for the Composition and Enactment of Bioinformatics Workflows." *Bioinformatics*, 20, 3045–3054.

Patra, A. K.,, et al. 2012. "A DDDAS Framework for Volcanic Ash Propagation and Hazard Analysis." In *Proceedings of the 2012 International Conference on Computational Science*. Elsevier.

Patrikalakis, N. M., et al. 2004. *Towards a Dynamic Data Driven Systems for Rapid Adaptive Interdisciplinary Ocean Forecasting*. http://robinson.seas.harvard.edu/PAPERS/kluwer_patrik_etal.pdf

Perumalla, K. S. 2007. "Scaling Time Warp-Based Discrete Event Execution to 10**4 Processors on a Blue Gene Supercomputer." In *Proceedings of the ACM Computing Frontiers Conference*. Ischia, Italy.

Pienta, R. and R. M. Fujimoto. 2013. "On the Parallel Simulation of Scale-Free Networks." *Principles of Advanced and Discrete Simulation*.

Riley, G. F., M. Ammar, R. M. Fujimoto, A. Park, K. S. Perumalla and D. Xu. 2004. "A Federated Approach to Distributed Network Simulation." *ACM Transactions on Modeling and Computer Simulation* 14(2): 116-148.

Scalable Network Technologies. 2012. "QualNet." 2012, from http://www.scalable-networks.com/content/products/qualnet.

Siganos, G., M. Faloutsos, P. Faloutsos and C. Faloutsos. 2003. "Power Laws and the AS-Level Internet Topology." *IEEE/ACM Transactions on Networking* 11(4): 514-524.

Soramaki, K., M. L. Bech, J. Arnold, R. J. Glass and W. E. Beyeler. 2007. "The Topology of Interbank Payment Flows." *Physica A: Statistical Mechanics and its Applications* 379(1): 317-333.

Stodden, V. 2010. “The Scientific Method in Practice: Reproducibility in the Computational Sciences.” Technical Report, MIT Sloan Research, February 2010.

Strassburger, S., T. Schulze and R. M. Fujimoto. 2008. “Future Trends in Distributed Simulation and Distributed Virtual Environments – Results of a Peer Study.” In *Proceedings of the 2008 Winter Simulation Conference*, 777-785. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Taylor, S. J. E., N. Mustafee, S. J. Turner and K. Pan. 2009. “Commercial-Off-The-Shelf Simulation Package Ineroperability: Issues and Futures.” In *Proceedings of the 2009 Winter Simulation Conference*, 203-215 editef by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin and R. G. Ingalls.

Walker, E. 2008. "Benchmarking Amazon EC2 for High Performance Scientific Computing." from http://www.usenix.org/publications/login/2008-10/openpdfs/walker.pdf.

Wang, X. F. and G. Chen. 2003. "Complex Networks: Small-World, Scale-Free and Beyond." *IEEE Circuits and Systems Magazine* 3(1): 6-20.

Wilson, A. L. and R. M. Weatherly. 1994. “The Aggregate Level Simulation Protocol: An Evolving System.” In *Proceedings of the 1994 Winter Simulation Conference*, 781-787.

Yilmaz, L. 2013. “Reproducibility in M&S Research: Issues, Strategies, and Implications for Model Development Environments.” *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 24, no. 4, pp. 457-474.

Yilmaz L. and J. Smith. 2008. “Prelude to the Panel on What Makes Good Research in Modeling and Simulation.” In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 671-677. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Zhang, L., X. Deng, J. Yu and X. Wu. 2011. "The Degree and Connectivity of Internet's Scale-Free Topology." *Chinese Physics B* 20(4): 048902.

AUTHOR BIOGRAPHIES

LEVENT YILMAZ is Professor of Computer Science and Software Engineering at Auburn University with a joint appointment in Industrial and Systems Engineering. He holds M.S. and Ph.D. degrees in Computer Science from Virginia Tech. His research interests are Modeling & Simulation, Agent-Directed Simulation, and Complex Adaptive Systems. He is the founding organizer and General Chair of the Annual Agent-Directed Simulation Symposium series and is currently serving as the Editor-in-Chief of the Simulation: Transactions of the SCS. His email address is yilmaz@auburn.edu.

SIMON J E TAYLOR is the Founder and Chair of the COTS Simulation Package Interoperability Standards Group (CSPI PDG) under SISO. He is the Editor-in-Chief of the Journal of Simulation. He was Chair of ACM SIGSIM from 2005 to 2008. He is a Reader in the Department of Computer Science and leads the Modeling & Simulation Group (http://tinyurl.com/os98k8y). He has published over 150 articles in modeling & simulation and distributed computing. His recent work has focused on leading the development of Grand Challenges in Modeling & Simulation, the development of standards for distributed simulation and cloud computing in industry, and advancing M&S in healthcare. His email address is simon.taylor@brunel.ac.uk.

RICHARD FUJIMOTO is Regents’ Professor in the School of Computational Science and Engineering at the Georgia Institute of Technology. He received a Ph.D. in Computer Science and Electrical Engineering from the University of California-Berkeley in 1983. His email address is fujimoto@cc.gatech.edu.

FREDERICA DAREMA is with the Air Force Office of Scientific Research (AFOSR). Prior to that, her career spans executive level positions at NSF for over 15 years, including detail assignment to DARPA as Program Manager, and before that she was for over 12 years as Research Staff Member and Research Manager at the IBM T. J. Watson Research Center, where she also served in the IBM Corporate Technical Strategy. She received her Ph. D. degree in Nuclear Physics from the University of California at Davis, where she attended as a Fulbright Scholar. Dr. Darema is Fellow of the IEEE. Her email address is frederica.darema@us.af.mil.