Introduction to the ICCS2006 Workshop on Dynamic Data Driven Applications Systems

Frederica Darema

National Science Foundation, Arlington VA 22230, USA
darema@nsf.gov

Abstract. The Dynamic Data Driven Application Systems (DDDAS) concept entails the ability to incorporate dynamically data into an executing application simulation, and in reverse, the ability of applications to dynamically steer measurement processes. Such dynamic data inputs can be acquired in real-time on-line or they can be archival data. DDDAS offers the promise of improving modeling methods, augmenting the analysis and prediction capabilities of application simulations, improving the efficiency of simulations and the effectiveness of measurement systems.

The scope of the present workshop provides examples of research and technology advances in enabling the DDDAS capabilities.

1 Introduction

The Dynamic Data Driven Application Systems (DDDAS) [1,2] concept entails the ability to incorporate dynamically data into an executing application simulation, and in reverse, the ability of applications to dynamically steer measurement processes. Such dynamic data inputs can be acquired in real-time on-line or they can be archival data. The DDDAS concept offers the promise of improving modeling methods, augmenting the analysis and prediction capabilities of application simulations, improving the efficiency of simulations and the effectiveness of measurement systems. Advances and technology capabilities required and enabled through the DDDAS concept are discussed in [3 and 4], and are fostered through the DDDAS Program [5] announced in 2005, with seeding efforts in the area having started previously (2000 – 2005) through the NSF ITR Program [6].   The DDDAS Program was co-sponsored by multiple Directorates and Offices of NSF, NOAA and NIH, and in cooperation with Programs in the European Community and the United Kingdom [referenced in 5].

In [4 and 5] it is emphasized the need for synergistic multidisciplinary research in applications modeling, in mathematical and statistical algorithms, in measurement methods and data management, and in computer systems software that's necessary to enable DDDAS capabilities. Furthermore in [4] the case is made that the DDDAS concept requires and will push for infrastructure that goes beyond the traditional notions of computational grids and provide new directions and capabilities in CyberInfractrucure environments [7]. While fundamental research advances are imperative to enable DDDAS capabilities, and the impact of DDDAS to scientific and

engineering fields is becoming widely recognized, there is increasing acceptance of the DDDAS potential to many industrial environments and such opportunities are manifested through the participation of industry in a number of the research projects.

The efforts presented in this workshop provide examples of technologies and capabilities that are being developed based on the DDDAS concept and supporting DDDAS environments. In www.cise.nsf.gov/dddas, a listing of all projects funded under the DDDAS rubric can be found; the papers presented here represent a sample of the range of projects and scope of the technical areas pursued. This introduction provides an overview and the context of the work presented. The present workshop is the fourth of a series of workshops on the DDDAS topic that have been organized as part of ICCS, starting with ICCS'03. Other community outreach activities include the NSF sponsored workshops [1 and 8].

2   Overview of Work Presented in This Workshop

The papers presented in this workshop represent ongoing research projects each addressing a number of aspects along the technical opportunities derived from the DDDAS concept as well as the challenges and technical advances required to enable DDDAS capabilities. With the exception of projects in [21] and [32-35], all other projects are enabled with the support of the ITR and DDDAS programs referenced above.

Several of the papers presented address application simulations whose capabilities are enhanced by dynamic data inputs into the executing application/simulation. Many of the papers presented involve multidisciplinary work to develop new capabilities along the four components: applications and algorithms, together with advances in computer systems software and measurement methods. While the emphasis of each project on each of the four components may vary, in most of these papers the application advances and enablement are made in synergism with in the underlying software technology, supporting the development of the complex DDDAS applications, their dynamic runtime support requirements and interfaces to control measurement systems. Several papers provide examples of simulation controlled measurement processes (instruments and sensor networks). Most of the work on mathematical algorithms tolerant and stable to dynamically streamed data is done in the context and together with the development of the application models. In all cases the mathematical algorithms and systems software development is done in the context of specific applications driving and validating the developed technologies.

The projects discussed in [9-19 and 22-30] enable improved analysis and prediction capabilities of application simulations and/or speeding up the computation times together with improved quality of simulations, and in a feedback loop the simulations control of measurements, sensors and other instruments.

In projects [9-12], the applications are in the areas of environmental and natural resource management. In [9] the project employs previous work performed by the investigators in the Instrumented Oil-Field DDDAS project that has enabled a new generation of data-driven, interactive and dynamically adaptive strategies for subsurface characterization and oil reservoir management. This previous work has led to the implementation of advanced multi-physics, multi-scale, and multi-block

numerical models and an autonomic software stack for DDDAS applications. The software stack implements a Grid-based adaptive execution engine, distributed data management services for real-time data access, exploration, and coupling, and self-managing middleware services for seamless discovery and composition of components, services, and data on the Grid. This paper investigates how these solutions can be leveraged and applied to address another DDDAS application of strategic importance, e.g. the data-driven management of Ruby Gulch Waste Repository. Work conducted in [10] develops capabilities to convert a traditional data collection sensor and ocean model into a DDDAS enabled system for identifying contaminants, dynamically coupling different models, simulations, and sensing strategies, in a symbiotic manner. The paper in [11] is centered on methods for detection, management and mitigation of threat in drinking water distribution systems; it involves real-time characterization of any contaminant source and plume, design of control strategies, and design of incremental data sampling schedules. This requires dynamic integration of time-varying measurements along with analytical modules hat include simulation models, adaptive sampling procedures, and optimization methods. These modules are compute-intensive, requiring multi-level parallel processing via computer clusters. Since real-time responses are critical, the computational needs must also be adaptively matched with available resources. This requires a software system to facilitate this integration via a high-performance computing architecture such that the measurement system, the analytical modules and the computing resources can mutually adapt and steer each other. The paper describes the development of such an adaptive cyberinfrastructure system facilitated by a dynamic workflow design. The paper in [12] is aimed at understanding environmental change through the deployment of next-generation wireless sensor networks and assimilating heterogeneous data, assessing the relative value and costs of data collection, and scheduling activities accordingly. Thus, they are dynamic, data-driven distributed systems that integrate sensing with modeling and prediction in an adaptive framework. Integration of a range of technologies will allow estimation of the value of future data in terms of its contribution to understanding and cost. This balance is especially important for environmental data, where sampling intervals will range from meters to entire landscapes, and from seconds to years in duration. The paper describes a general framework for dynamic data-driven wireless network control that combines modeling of the sensor network and its embedding environment, both in and out of the network, and describes a range of challenges that must be addressed, and an integrated suite of solutions for the design of dynamic sensor networks.

The papers in [13-15] are in application areas dealing with security and crisis management and mitigation. The paper in [13] is developing a prototype emergency response system. This dynamic data driven application system (DDDAS) uses wireless call data, including call volume, who calls whom, call duration, services in use, and cell phone location in formation to utilize the cell phones of a city to serve as an ad hoc mobile sensor net, measuring the movement and calling patterns of the population. Social network theory and statistical analysis on normal call activity and call locations establish a baseline. A detection and alert system monitors streaming summary cell phone call data. Abnormal call patterns or population movements trigger a simulation and prediction system. The project in [14] is concerned with applying Dynamic Data Driven Application Simulations (DDDAS) to monitor and

manage surface transportation systems. Building upon activities such as the Vehicle-Infrastructure Integration initiative, a hierarchical DDDAS architecture will be created that includes coupled in-vehicle, roadside, and traffic management center simulations. The overall architecture and further advances incorporating the DDDAS concept will be deployed to implement and evaluate the effectiveness of the intended system for a portion the Atlanta area in the context of a hypothesized emergency evacuation scenario. The paper in [15] will apply DDDAS capabilities developed by the investigators that allow to study interaction between fire and agent models during a fire evacuation. The analysis from that research can be used to validate proposed ideas in evacuation and building designs to ensure safety of buildings given various agent behaviors. Through study of the interactions, a better understanding is gained of how to improve building design, evacuation strategies and regulations, as well as training first responders.

The papers in [16-17] are addressing application areas of critical infrastructure management and protection, such as electrical power systems management   The paper in [16] is focused in developing methods and approaches for maintaining acceptable electric transmission system reliability and optimized electric energy delivery, requires innovations in sensing, diagnostics, communications, data management, processing, algorithms, risk assessment, decision-making (for operations, maintenance, and planning), and process coordination. The work presented is a comprehensive approach to develop methods and processes in these areas, driven by the ultimate objective to develop a hardware-software prototype capable of auto-steering the information-decision cycles inherent to managing operations, maintenance, and planning of the high-voltage electric power transmission systems. The paper in [17] is also concerned with fault-tolerant operation of large-scale electric power systems through efficient and methods that employ the DDDAS concept to enable fast and robust reduced-order-modeling and filtering methods and merging of data from multiple sources.

The papers in [18-21] address methods for structural monitoring and fault-tolerance in structures like buildings and aircraft. Paper [18] discusses methods and implementation scope supporting a dynamic data driven application system for material and structural health monitoring as well as critical event prediction. The dynamic data driven paradigm is exploited to promote application advances, application measurement systems and methods, mathematical and statistical algorithms and systems software infrastructure relevant to this effort. These advances are intended to enable behavior monitoring and prediction as well as critical event avoidance on multiple time scales. Paper [19] uses DDDAS to enable structural sensing and control. Paper [20] is employing DDDAS for evaluation of fluid-thermal systems wherein a complete specification of the boundary conditions is not known a priori and experimental diagnostics are restricted to a limited region of the flow-field. The methodology is applied to the configuration of a heated jet injected into a laminar boundary layer where the jet temperature is not known a priori. Papers [21, 37] use DDDAS to improve methods for analysis, prediction of atmospheric contaminant propagation and inversion for identification of the origin of the contaminant, employing the concept to steer sensors and stream up-to-date information to enhance reduced order models and PDE solvers approaches, in a feed-back control loop.

Papers [22-26] address topics of atmospheric, aqueous and ambient environmental systems such as: In [22] a data assimilation technique for coupled ionospheric and thermospheric dynamics. Using a frozen linear dynamics matrix for the time update of the error covariance and the evaluation of the Kalman filter gain, and demonstrating the performance of the data assimilation technique on a section of the ionosphere. Paper [23] uses the DDDAS concept to simultaneously factor position and amplitude errors for improved prediction of path of rapidly changing mesoscale events like hurricanes. In [24], in the context of a marine biology application, observing systems facilitate scientific studies by instrumenting the real world and collecting corresponding measurements, with the aim of detecting and tracking phenomena of interest. Thos project focuses on a class of observing systems, which are embedded into the environment, consist of stationary and mobile sensors, and react to collected observations by reconfiguring the system and adapting which observations are collected next. Paper [26] reports on an ongoing effort to build a Dynamic Data Driven Application System (DDDAS) for short-range forecast of weather and wildfire behavior from real-time weather data, images, and sensor streams. The system changes the forecast as new data is received. The methods involve applying an ensemble Kalman filter in time-space with a highly parallel implementation, and demonstration of the validity of the derived system by using a DDDAS test-bed approach and data collected from an earlier fire.

Applications in [27-29] are in the areas of health and behavioral sciences. The project in [27] uses the DDDAS concept to enable a dynamic data-driven planning and control system for laser treatment of cancer. The research includes development of a general mathematical framework and a family of mathematical and computational models of bio-heat transfer, tissue damage, and tumor viability, dynamic calibration, verification and validation processes based on laboratory and clinical data and simulated response, and (3) design of effective thermo-therapeutic protocols using model predictions. At the core of the proposed systems is the adaptive-feedback control of mathematical and computational models based on a posteriori estimates of errors in key quantities of interest, and modern Magnetic Resonance Temperature Imaging (MRTI), and diode laser devices to monitor treatment of tumors in laboratory animals. This approach enables an automated systematic model selection process based on acceptance criteria determined a priori. The methodologies to be implemented involve uncertainty quantification methods designed to provide an innovative, data-driven, patient-specific approach to effective cancer treatment. Paper [28] presents a case of interdisciplinary collaboration in building and using a set of tools to compute the flows in a branched artery, to compare them with prior physical flow visualization, and to interpret them with further users in mind. The geometry is of a typical epicardial coronary artery with a side branch. The incompressible Navier-Stokes equations are solved with the hybrid spectral/hp element solver Nektar, and simulations were visualized in the CAVE, an immersive 3D stereo display environment, which facilitates interpretation of simulated features. Paper [29] describes a novel approach towards the development of new design architectures and research test-beds for advanced Brain Machine Interfaces (BMIs) to help paralyzed patients, and others with motor disabilities, regain (artificial) motor control and autonomy. It addresses a critical design challenge in deriving the functional mapping between the subject's movement intent and actuated

behavior. Inspired by motor control research, this paper considers a predictive framework for BMI using multiple adaptive models trained with supervised or reinforcement learning in a closed-loop architecture that requires real-time feedback. In this framework, BMIs require a computing infrastructure capable of selectively executing multiple models on the basis of signals received by and/or provided to the brain in real time. Middleware currently under investigation to provide this data-driven dynamic capability is discussed. Paper [30] presents a dynamic data-driven framework for tracking gestures and facial expressions from monocular sequences. The system uses two cameras, one for the face and one for the whole body view for processing in different scales. Specifically, and for the gesture tracking module, the system tracks the hands and the head, obtaining as output the blobs (ellipses) of the ROIs, and detects the shoulder positions with straight lines. For the facial expressions, the approach is to extract the 2*D* facial features, using a fusion between KLT tracker and a modified Active Shape Model, and then we obtain the 3D facemask with fitting a generic model to the extracted 2*D* features. The main advantages of the system pursued are: adaptivity, i.e. robustness to external conditions, such as lighting, and independent from the examined individual, and computational efficiency, providing us results off- and online with a rates higher than 20*frames/sec.*

In [31] the work presented employs artificial intelligence (AI) approaches to enable the management of a data driven simulation system, in particular with regard to adaptive selection of data and refinement of the model on which the simulation is based. The approaches used consider two different classes of intelligent agent that can control a data driven simulation: (a) an autonomous agent using internal simulation to test and refine a model of its environment and (b) an assistant agent managing a data-driven simulation to help humans understand a complex system (assisted model-building). In the first case the agent is situated in its environment and can use its own sensors to explore the data sources. In the second case, the agent has much less independent access to data and may have limited capability to refine the model on which the simulation is based. This is particularly true if the data contains subjective statements about the human view of the world, such as in the social sciences. For complex systems involving human actors, the paper pursues an architecture in which assistant agents cooperate with autonomous agents to build a more complete and reliable picture of the observed system.

The paper in [32] presents tools that serve to capture and apply a scientist's insight about opportunities for, and limitations of, simulation adaptation, and reports on the two ongoing collaborations that serve to guide and evaluate the methods pursued. One of the intended consequences of utilizing simulations in dynamic, data-driven application systems is that the simulations will adjust to new data as it arrives. These adjustments will be difficult because of the unpredictable nature of the world and because simulations are so carefully tuned to model specific operating conditions. Accommodating new data may require adapting or replacing numerical methods, simulation parameters, or the analytical scientific models from which the simulation is derived. This research project emphasizes the important role a scientist's insight can play in facilitating the runtime adaptation of a simulation to accurately utilize new data.

The papers [33-36] present work that can be evolved as enabling algorithmic and computer systems software technologies supporting DDDAS environments. In [33] web services, service oriented architecture and model driven architecture approaches

are discussed; in [34] advanced visualization methods, in [35] multivariate signal analysis methods are discussed, and in [36] knowledge-based methods for data management approaches are presented. The methods discussed in these papers are needed advances for capabilities such as dynamic composition of DDDAS applications and dynamic data management interfaces.

3 Summary

The DDDAS concept opens new capabilities in application and measurement methods and systems. Through efforts that have started developing DDDAS systems, such promise begins to materialize, and we are in a steady path towards accomplishing that objective. The progress made is exemplified by the projects presented here.

Acknowledgements

References

1. NSF Workshop, March 2000; www.cise.nsf.gov/dddas
2. F. Darema, Dynamic Data Driven Applications Systems: A New Paradigm for Application Simulations and Measurements, ICCS'04
3. F. Darema, Grid Computing and Beyond: The Context of Dynamic Data Driven Applications Systems, Proceedings of the IEEE, Special Issue on Grid Computing, 3/05
4. F. Darema, Dynamic Data Driven Applications Systems: New Capabilities for Application Simulations and Measurements, ICCS'05
5. DDDAS-Dynamic Data Driven Applications Systems Program Solicitation (NSF 05-570); www.cise.nsf.gov/dddas
6. NSF Information Technology Research (ITR) Program (1999-2004)
7. Cyberinfrastructure Report http://www.communitytechnology.org/nsf_ci_report
8. NSF Sponsored Workshop: DDDAS-Dynamic Data Driven Applications Systems, Jan 19-20, 2006 [www.cise.nsf.gov/dddas]
9. Manish Parashar, Vincent Matossian, Hector Klie, Sunil G. Thomas, Mary F. Wheeler, Tahsin Kurc, Joel Saltz, and Roelof Versteg; Towards Dynamic Data-Driven Management of the Ruby Gulch Waste Repository
10. Craig C. Douglas, J. Clay Harris, Mohamed Iskandarani, Chris R. Johnson, Robert Lodder, Steve Parker, Martin J. Cole, Richard Ewing, Yalchin Efendiev, Raytcho Lazarov, and Guan Qin; Dynamic Contaminant Identification in Water
11. Kumar Mahinthakumar, Gregor von Laszewski, Ranji Ranjithan, Downey Brill, Jim Uber, Ken Harrison, Sarat Sreepathi, and Emily Zechman; An Adaptive Cyberinfrastructure for Threat Management in Urban Water Distribution Systems:
12. Paul G. Flikkema, Pankaj K. Agarwal, James S. Clark, Carla Ellis, Alan Gelfand, Kamesh Munagala, and Jun Yang; Model-Driven Dynamic Control of Embedded Wireless Sensor Networks

13. Greg R. Madey, Gabor Szabo, and Albert-László Barabási; WIPER: The Integrated Wireless Phone Based Emergency Response System
14. Richard M. Fujimoto, Randall Guensler, Michael Hunter, Hoe Kyoung Kim, Jaesup Lee, John Leonard II, Mahesh Palekar, Karsten Schwan, Balasubramanian Seshasayee; Dynamic Data Driven Application Simulation of Surface Transportation Systems
15. Alok Chaturvedi, Angela Mellema, Sergei Filatyev, Jay Gore; DDDAS Approach to Fire and Agent Evacuation Modeling: Case Study of Rhode Island Nightclub Fire
16. James D McCalley, Vasant G Honavar, Sarah M Ryan, William Q Meeker, Ronald A Roberts, Daji Qiao, Yuan Li; Auto- Steered Information-Decision Processes for Electric System Asset Management
17. E. H. Abed, N. S. Nmachchivaya, T. J. Overbye, M. A. Pai, P. W. Sauer and A. Sussman; Data-Driven Power System Operations
18. C. Farhat, J. G. Michopoulos, F.K. Chang, L.J. Guibas, and A.J. Lew; Towards a Dynamic Data Driven System for Structural and Material Health Monitoring
19. Asad Awan, Ahmed Sameh, and Ananth Grama; The Omni Macroprogramming Environment for Sensor Networks
20. D. Knight, T. Rossman and Y. Jaluria; Evaluation of Fluid-Thermal Systems by Dynamic Data Driven Application Systems
21. Volkan Akcelik, George Biros, Andrei Draganescu, Omar Ghattas, Judith Hill, and Bart van BloemenWaanders; Inversion of Airborne Contaminants in a Regional Model
22. S. Kim, J. Chandrasekar, A. Ridley, D. S. Bernstein; Data Assimilation Using the Global Ionosphere-Thermosphere Model
23. Sai Ravela; Amplitude-Position formulation of Data Assimilation
24. Hyung-Jin Son and Theodore B. Trafalis; Detection of Tornados Using an Incremental Revised Support Vector Machine with Filters
25. Leana Golubchik, David Caron, Abhimanyu Das, Amit Dhariwal, Ramesh Govindan, David Kempe, Carl Oberg, Abhishek Sharma, Beth Stauer, Gaurav Sukhatme, Bin Zhang; A Generic Multi-scale Modeling Framework for Reactive Observing Systems: an Overview
26. Craig C. Douglas, Jonathan D. Beezley, Janice Coen, Deng Li, Wei Li, Alan K. Mandel, Jan Mandel , Guan Qin, and Anthony Vodacek; Demonstrating the Validity of a Wildfire DDDAS
27. J. T. Oden, K. R. Diller, C. Bajaj, J. C. Browne, J. Hazle2, I. Babu¡ska, J. Bass, L. Demkowicz, Y. Feng, D. Fuentes, S. Prudhomme, N. Rylander,R. J. Sta_ord, and Y. Zhang; Development of a Computational Paradigm for Laser Treatment of Cancer
28. P. D. Richardson, I.V. Pivkin, G.E. Karniadakis, D.H. Laidlaw; Blood Flow At Arterial Branches: Complexities To Resolve For The Angioplasty Suite
29. José Fortes, Renato Figueiredo, Linda Hermer-Vazquez, José Príncipe and Justin C. Sanchez; A New Architecture for Deriving Dynamic Brain-Machine Interfaces
30. D. Metaxas, G. Tsechpenakis, Z. Li, Y. Huang, A. Kanaujia; Dynamically Adaptive Tracking of Gestures and Facial Expressions
31. Catriola Kennedy and Georgios Theodoropoulos;: Intelligent Management of Data Driven Simulations to Support Model building in the Social sciences
32. Paul Reynolds, David Brogan, Joseph Carnahan, Yannick Loiti`ere, and Michael Spiegel; Capturing Scientists' Insight for DDDAS
33. Adel Torkaman Rahmani, Vahid Rafe, Saeed Sedighian and Amin Abbaspour; An MDA-based Modeling and Design of Service-Oriented Architecture
34. Anthony Jones and Dan Cornford; Advanced Data Driven Visualisation for Geo-spatial Data

35. Liu Bo, Zhao Jun, and Qian Jixin; Design and Analysis of Test Signals for System Identification
36. XiaoDong Gao and Zhiping Fan; The Research on the Method of Process-based Knowledge catalog & Storage and its Application in Steel Product R&D
37. Emil M. Constantinescu, Adrian Sandu, Gregory R. Carmichael, Tianfeng Chai, John H. Seinfeld, and Dacian D¢aescu; Localized Ensemble Kalman Data Assimilation for Atmospheric Chemical Transport Models