Dynamic Data Driven Application Systems

Overview:

This workshop will focus on recent advances in Dynamic Data-Driven Application Systems (DDDAS). DDDAS seeks to enable new capabilities through the synergistic integration of models and data in a dynamic feedback control loop, whereby instrumentation data of a system are dynamically integrated with simulation to enhance modeling and prediction, while models, in turn, can be used to enhance the effective use of system instrumentation, including networks of heterogeneous sensors and controllers. Application models can encompass large numeric and non-numeric methods as well as analytical methods. Representative areas where DDDAS applies range from environmental applications, such as weather forecasting, forest fire prediction, to space situational awareness, to critical infrastructure systems, such as transportation systems, electrical power grids and other energy systems, to medical applications, and more broadly for the analysis of natural and engineered systems, and adaptive management of complex systems, including distributed systems, with autonomic and fault-tolerant operation needs.

The aim of this workshop is to bring together researchers working in the DDDAS area and present some recent results to the ACC community under one unifying theme. This is expected to induce synergistic discussions and future interaction between the presenters and the researchers in the audience. This workshop will highlight recent advances in modeling and analysis, data collection and management, mathematical and statistical algorithms for uncertainty characterization and forecasting, data assimilation, and sensor planning.

At the end of this workshop, audience will:

1. Understand and use the concept of Dynamic Data Driven Application Systems.
2. Gain exposure to implementation issues such as computational complexity, uncertainty in model and data, data assimilation, and dynamic data monitoring.

Various academic and engineering problems where traditional methods either fail or perform very poorly will be considered to demonstrate the reliability and limitations of the newly established methods.

Format:

This workshop will consist of one half-day session. The focus of this session will be to introduce the basic concept of DDDAS, and application of various DDDAS methodologies to real engineering applications. Speakers will lead sessions by giving short, concise presentations on the latest developments, followed by a forum of speakers-audience interactions on questions and concerns from participants, submitted either on-line or in person. Practical problems and examples will be used throughout the discussions to help participants connect novel principles to their own experiences.

Target Audience:

Practitioners looking for advanced methods to propagate non-Gaussian state and parameter uncertainties through nonlinear dynamical systems, researchers looking for new ideas and the connection between theory and practice; students looking for better understanding of the foundation of uncertainty quantification and the essence of nonlinear estimation engineering practice.

Tentative Schedule and Speakers:

Speakers: Dr. Frederica Darema (Air-force Office of Scientific Research), Dr. Dennis Bernstein (University of Michigan, Ann-Arbor), Dr. Sai Ravela (Massachusetts Institute of Technology), Dr. Puneet Singla (SUNY Buffalo

SESSION 1. (8:00 – 12:00)

1) Introduction and Opening Remarks (8:00 – 8:15)
2) Why DDDAS (8:15 - 9:00)
3) Data-Based Techniques for Model Calibration, Correction, and Refinement (9:15 - 10:00)
4) Learning, Information and Perception in Environmental Systems Science (10:15 – 11:00)
5) Dynamic Data Monitoring for Tracking Resident Space Objects and Atmospheric Release (11:15 – 12:00)

Abstracts for Talks:

Why DDDAS: (Speaker: Dr. Frederica Darema) – This talk will introduce the basic concept of Dynamic Data Driven Application Systems (DDDAS). DDDAS embodies the power of Dynamic Data Driven Applications Systems (DDDAS), a concept whereby an executing application model is dynamically integrated, in a feed-back loop, with the real-time data-acquisition and control components, as well as other data sources of the application system. Advanced capabilities can be created through such new computational approaches in modeling and simulations, and in instrumentation methods, and include: enhancing the accuracy of the application model; speeding-up the computation to either allow more comprehensive models of a system, or create decision support systems with the accuracy of full-scale simulations. In addition, the notion of controlling instrumentation processes by the executing application results in more efficient management of application data and addresses challenges of how to architect and dynamically manage large sets of heterogeneous sensors and controllers, an advance over the static and ad-hoc ways of today - with DDDAS these sets of resources can be managed adaptively and in optimized ways. Supporting the integrated computational and instrumentation aspects of InfoSymbiotics/DDDAS environments entails a unified computational-instrumentation underlying platform.

The talk will address opportunities for such new capabilities together with corresponding research challenges in applications, application mathematical and statistical algorithms, instrumentation, and systems software, with illustrative examples from several application areas where InfoSymbiotics/DDDAS engenders transformative impact; examples include: analysis and decision support for structural engineered systems, environmental systems, critical infrastructure systems, such as electrical powergrids, and manufacturing systems and processes.

Data-Based Techniques for Model Calibration, Correction, and Refinement (Speaker: Dr. Dennis Bernstein) – It is often said that "all models are wrong, but some are useful." With this utilitarian view in mind, we recognize that most models are based on a combination of physics and data, where the physics provide the backbone of the model and data are used to fill in details

that are difficult to specify from physics alone. This portion of the workshop will focus on the following question: How can data be used to improve the accuracy of a given model? This is a problem in model calibration, correction, and refinement, where the goal is to use data to modify the original model so that it more accurately represents the underlying physical system. This problem is relevant to a wide range of applications that require accurate models, such as large-scale data assimilation for DDDAS.
Model refinement can be viewed as a specialized problem in system identification, where, unlike the usual case, where a complete model is identified, the goal is to identify a subsystem of a larger model. The subsystem may be static or dynamic, and its inputs and outputs may be unmeasured. In this session, model refinement will be addressed through retrospective cost techniques. The theory and algorithm will be reviewed, and applications will include illustrative examples from circuits and structures, as well as large-scale applications such as the identification of missing physics in a global model of the ionosphere. Finally, the same technique will be used to reconstruct unknown drivers for use in large-scale data assimilation.

Learning, Information and Perception in Environmental Systems Science (Speaker: Dr. Sai Ravela) – Dynamic data-driven systems science is valuable for environmental investigations. Non-linear processes represented by high-dimensional numerical models with uncertain parameters, states, forcing and structure require data in order to be effective tools for analysis , prediction and discovery. The coupling of models and measurements takes on a variety of forms that require inferences from data to understand models, inferences from data and models to understand the physical system state, and inferences from models to reason about data. Several inference problems can be expressed in terms of fixed-interval, fixed-lag or fixed-point estimation problems where the need is efficacy in the presence of high-dimensional, non-Gaussian state-spaces with sparse, noisy measurements. In this talk, using state estimation as a prototypical inference problem, we begin with a review of two-point boundary value problems and sequential Bayesian state estimation culminating in the particle filter and smoother. We then discuss three directions of interest.

Learning: Error distributions of model predictions are poorly understood. Even, for example, as one must deal with emerging non-Gaussian uncertainties in nonlinear system states, overall variance remains an important objective in the presence of model error and cannot be ignored. We approach inference using multiple distributions to model the underlying uncertainty, and propose an Ensemble Learning framework for non-Gaussian estimation. We show how this framework produces estimators with tighter posterior uncertainties and robustness to model bias.

*Information***:** Sampling remains a critical problem in non-Gaussian inference. One way to ameliorate this problem is to approach non-Gaussian estimation in part as an optimization problem. Using the mutual information as a "distribution free" summary measure is relatively well recognized but does not by itself reduce sampling burdens. Less known and more interesting is to employ quadratic forms of mutual information which we argue leads to tractable non-Gaussian estimation via gradient-based optimization. After a review of this framework, we show that quadratic mutual information can be effectively optimized for high dimensional non-Gaussian distributions. In particular, we examine techniques ranging from fast gauss transform to hashing and show that a slight relaxation of exactness can lead to highly efficient and effective non-parametric estimators. We additionally propose a scale-space information graph for added efficiency and show how model identification of this graph is cost effective. With these techniques in hand, we will build filters, fixed-lag and fixed-interval smoothers that are both adjoint free and require minimal sampling.

Perception: The analysis, synthesis and representation of coherent fields is fundamental to many spatial inference problems. We first show the failure of classical methods in the context of coherent fields and then synthesize an effective approach utilizing the above results in the context of a pattern theory for fluids. In this theory, we posit that position (deformation, geometry) and amplitude (appearance, intensity) both play a significant role, which leads to several new approaches. These include super-resolution, reduced models, targeting (adaptive sampling), data assimilation (state estimation), and uncertainty quantification.

Dynamic Data Monitoring for Tracking Resident Space Objects and Atmospheric Release (Speaker: Dr. Puneet Singla) – The main focus of this talk is to introduce a dynamic data monitoring framework for the estimation of spatio-temporal dynamical variables by optimally managing mobile sensors, which is conducive to scaling to accommodate increasing numbers of sensors. The crux of the work lie in accounting for uncertainties in model parameters of systems driven by stochastic input, characterizing the evolution of the uncertainty of the forecast variables, and integrating disparate sources of asynchronous data with the model output using a Bayesian framework. Intimately tied to the data assimilation problem is the optimal deployment of mobile sensors (e.g., UAV) with the objective of minimizing the uncertainty of the forecast variable, which corresponds to the maximization of the degree of Situational Awareness (SA) so as to facilitate decision making. The most critical challenge here is to provide a quantitative assessment of how closely our estimates reflect reality in the presence of model uncertainty, discretization errors as well as measurement errors and uncertainty. The quantitative understanding of uncertainty is essential when predictions are to be used to inform policy making or mitigation solutions where significant resources are at stake.

This talk will focus on recent development of mathematical and algorithmic fundamentals for uncertainty propagation, forecasting, model-data fusion and optimal control for nonlinear systems. The central idea is to replace evolution of initial conditions for a dynamical system by evolution of probability density functions (pdf) for state variables. Various academic and engineering problems such as tracking resident space objects and forecasting of toxic material plume being advected through the atmosphere, where traditional methods either fail or perform very poorly, are considered to assess the reliability and limitations of the newly established methods. Some results from these studies will be discussed.

Bio-sketches for Speakers:

Frederica Darema: Dr. Frederica Darema, a member of the Senior Executive Service, is the Director, Mathematics, Information and Life Sciences, Air Force Office of Scientific Research, Arlington, Va. She provides executive direction in the planning, conduct and coordination of broad, frequently large-scale, and critical basic research and development program activities. These include the areas as advanced mathematical and computational methods for dynamic systems; information and decision systems; bio-systems; human cognition and socio-cultural systems.

Dr. Darema is a graduate of the University of Athens, Greece; the Illinois Institute of Technology; and the University of California at Davis, where she attended as a Fulbright Scholar and a Distinguished Scholar. After physics research associate positions at the University of Pittsburgh and Brookhaven National Laboratory, she received an American Physics Society Industrial Postdoctoral Fellowship and became a technical staff member in the Nuclear Sciences Department at Schlumberger-Doll Research. Subsequently, she joined the T.J. Watson IBM Research Center as a research staff member and group manager. While at IBM, she also served in the IBM Corporate Strategy Group examining and helping to set corporate-wide strategies. From

1996 to 1998, she completed a two-year interagency assignment at the Defense Advanced Research Projects Agency.

Prior to her current assignment, Dr. Darema was at the National Science Foundation where she held executive level positions as Senior Science and Technology Adviser and Senior Science Analyst in the Computer and Information Science and Engineering Directorate. She has given numerous keynote speeches and presentations in professional forums. Dr. Darema's scientific and technical contributions include development of parallel applications; parallel algorithms; programming models; environments and systems performance engineering for the design of applications and of software for parallel and distributed systems.

Dennis Bernstein: Dr. Dennis Bernstein is a professor in the Aerospace Engineering Department of the University of Michigan. He is a former Editor-in-Chief of the IEEE Control Systems Magazine, an IEEE Fellow, and author of Matrix Mathematics (2nd ed., 2009).

Sai Ravela: Dr. Sai Ravela directs the Earth Signals and Systems Group in the Earth, Atmospheric and Planetary Sciences department at MIT. His interdisciplinary group develops theory, methodology and experimental approaches to environmental systems science. Current projects include animal biometrics (sloop.mit.edu), perceptual representations of coherent fluids (stics.mit.edu), cooperative autonomous environmental observation (caos.mit.edu), extreme event risk and uncertainty prediction (hurriup.mit.edu), and fluid imaging and sensing (flux.mit.edu). Sai Ravela received a PhD in 2002 in Computer Science with specialization in Artificial Intelligence from the University of Massachusetts at Amherst, and post-doctoral training in Stochastic Systems and Atmospheric Science at MIT. He is co-founder of Windrisktech LLC, a hurricane risk prediction company using machine learning and hurricane physics, and E5 Aerospace LLC which develops cooperative autonomous sampling platforms. For more information visit htttp://essg.mit.edu.

Puneet Singla: Dr. Puneet Singla is an Associate Professor of Mechanical & Aerospace engineering at the University at Buffalo (UB), the State University of New York. He received his bachelor’s degree in Aerospace Engineering from Indian Institute of Technology, Kanpur, India in 2000 and earned his master’s and doctoral degree in Aerospace Engineering from Texas A&M University, College Station in 2002 and 2006, respectively. His research work includes three thrusts: 1) characterization and propagation of uncertainties through dynamical systems, 2) design of computationally efficient data assimilation algorithms for large scale problems, and 3) design robust methodologies for optimal sensor management while taking into account the uncertainties in the system dynamics.

During his tenure at UB, he has secured several research grants as a PI or co-PI from the National Science Foundation (NSF), Air Force Office of Scientific Research (AFOSR), Air Force Research Laboratory (AFRL) and the National Geospatial Intelligence Agency (NGA). He is a recipient of the competitive NSF CAREER award for his work on Uncertainty Propagation and Data Assimilation for Toxic Cloud Prediction and the AFOSR Young Investigator Award for his work on Information Collection and Fusion for Space Situational Awareness. He has also been awarded the UB’s “Exceptional Scholar” Young Investigator Award in recognition of his scholarly activities. He has authored over 100 papers to-date including 25 journal articles covering a wide array of problems, including: attitude estimation, nonlinear estimation, dynamics and control, adaptive control, approximation theory, including novel methods for solving the Fokker-Planck-Kolmogorov equation (FPKE) for uncertainty propagation.