Posters at AD 2012

1. <u>Patrick E. Farrell</u> (Department of Earth Science and Engineering, Imperial College London, UK):
Automating the adjoint of finite element discretisations

In this work we demonstrate the capability of automatically deriving the associated discrete adjoint and tangent linear models from a forward model written in a high-level finite element computing environment.

High-level systems allow the developer to express the variational form to be solved in near-mathematical notation, from which the low-level code is automatically generated by a custom finite element compiler. By exploiting optimisations that are impractical to perform by hand, such systems can generate extremely efficient model implementations. These high-level systems have a further key advantage: because the mathematical structure of the problem to be solved is preserved, they are much more amenable to automated analysis and manipulation than a conventional model written in a low-level language (C/Fortran).

We present a software package capable of automatically deriving adjoint and tangent linear models from forward models written in the DOLFIN finite element computing environment. The approach relies on run-time annotation of the temporal structure of the model, and employs the same finite element form compiler to automatically generate the assembly code for the derived models. This method requires only trivial changes to most forward models, and works even for complex time-dependent nonlinear forward models.

Taking this high-level approach to adjoint model generation has several major advantages. The adjoint model automatically employs optimal checkpointing schemes to mitigate storage requirements for nonlinear models, without any user management or intervention. The approach extends even to models which use matrix-free solvers; differentiating through matrix-free solves is typically very difficult for a low-level automatic differentiation tool. Furthermore, both the tangent linear and adjoint models naturally work in parallel, without any need to differentiate through calls to MPI.

The efficiency and generality of the approach are demonstrated with examples from several different scientific applications.

2. <u>Koichi Kubota</u> (Chuo University, Tokyo, Japan):
Modular Arithmetic Approach to Combinatorial Computation derived from AD

Counting the number of Hamiltonian cycles of a given graph that is one of typical combinatorial computations can be formulated with higher order derivatives. Its value is computed as the residue with complex floating numbers. But there are inevitable rounding errors in the conventional computation of the residue whereas the mathematical result is an integer value. In this paper algorithms of the residue with modular arithmetics are proposed in order to compute the exact integer result. It is shown that they are naturally executed on parallel computers by partitioning the summation, and that their computational speed is almost proportional to the number of processors with numerical experiments using 100 machines.

3. <u>Maria Brune</u>, Andrea Walther (University of Paderborn, Germany);
Christoph Statz, Sebastian Hegler, Dirk Plettemeier (Technical University Dresden, Germany):
Numerical studies on inverse electromagnetic scattering problems using Python-based tools

One significant reason of undesirable losses of stationary technical stream machines is the frequent appearance of radial gaps between rotating turbine blades and the turbine wall. Common strategies cannot be applied to this kind of problems for several reasons, e.g. inappropriate geometric shapes of the object, heat or dirty lenses. Therefore, radar measurements can be applied where electromagnetic waves penetrate a certain domain. Consequently, the analysis of an inverse electromagnetic scattering problem is the focus of our work. The goal is to obtain the locations of certain geometric bodies in a predefined area. In previous work, we reconstructed the distribution of the dielectric permittivity of ellipsoids on a discretized grid for each grid cell. This led to a high number of unknowns such that the AD reverse mode was an appropriate choice. In our current work, we consider the reconstruction of the location of ellipsoids where the center and the radii in x, y and z direction have to be identified. Due to the small number of independent variables, the AD forward mode is a good choice here.

First, we placed an ellipsoid with a given conductivity in the computational domain. The forward simulation is performed by an appropriate FDTD-method, which is developed at TU Dresden in the context of the cooperation project HPC-FLiS. This method discretizes the resulting fields. A point of interest is to minimize the discrepancy between the observed and simulated data. The arising optimization problem is solved by the software package cyipopt, which provides bindings for the interior point method IPOPT. The required derivatives are calculated by pyadolc, developed by Sebastian F. Walter, a Python wrapper for the automatic differentiation tool ADOL-C. For our application, the calculation of the derivatives is a challenging task, since the right placement of the conductivity parameters in the computational domain depends on an indicator function. A suitable strategy to overcome this problem will be presented. From theory we obtain a proportionality between the computation time of one derivative computation and the time of one function evaluation. Corresponding scalability tests were performed and analyzed. Moreover, first approaches concerning parallelization by MPI have been evaluated and tested. Recent theoretical and numerical results will be presented and discussed.

4. <u>Sabrina Fiege</u> (University of Paderborn, Germany); Andreas Griewank (Humboldt University Berlin, Germany); Andrea Walther (University of Paderborn, Germany):
An exploratory line-search for piecewise smooth functions

Many scalar or vector functions occurring in numerical applications are not continuously differentiable. This situation arises in particular through the use of l_1 or l_∞ penalty terms or the occurrence of abs(), min() and max() in the problem function evaluations themselves. If no other nonsmooth elemental functions are present, generalized gradients and Jacobians of these piecewise smooth functions can now be computed in an AD like fashion by *lexicographic differentiation* as introduced by Barton & Kahn, Griewank and Nesterov. Therefore, we allow in contrast to the standard approach in algorithmic or automatic differentiation (AD) the occurrence of the mentioned nonsmooth but Lipschitz continuous functions during the function evaluation. This yields piecewise differentiable nonlinear functions, that can be approximated locally by piecewise-linear models. These models will have at most a finite number of kinks between open polyhedra decomposing the function domain. The piecewise linear model can easily be generated by minor modification of ADOL-C and other standard AD tools. However, the trouble is that at almost all points these generalized derivatives reduce to classical derivatives, so that the effort to provide procedures that can calculate generalized Jacobians almost never pay off. At the same time the Taylor approximations based on these classical derivative singeltons may have a miniscule range of validity. On this poster we show how along a given search direction one can either compute the critical step multiplier that leads to either the nearest kink or to the first local minimum or to the global minimum. Naturally, the achievement of these goals can not be guaranteed absolutely, but we verify necessary conditions based on cubic interpolation or bisection using one-sided derivatives. The line-search is working well for small-scale nonsmooth problems. For large-scale problems new strategies have to be developed. The poster also presents numerical results.

5. <u>David M. Gay</u> (AMPL Optimization, Inc., USA):
Bounds from Slopes

Sometimes it is desirable to compute good bounds on the possible values of an algebraic expression that involves variables only known to lie in prescribed finite intervals. While interval arithmetic gives rigorous bounds, if some variables appear more than once in the expression, bounds from interval arithmetic can be very pessimistic. Propagation of Taylor series has long been known as a way to obtain much tighter bounds. More recently, researchers have shown that "slope" computations often give tighter bounds than Taylor series. First-order slope computations are fairly straightforward and are closely akin to forward AD, but second-order slope computations (which also involve forward AD) sometimes give still better bounds. This presentation reviews interval, Taylor-series and slope computations, presents a possibly new bound optimization, and discusses implementation approaches based on expression graphs and operator overloading. It also describes a still unfinished implementation for experimenting with expression graphs. The implementation should eventually be available as open source in the AMPL/solver interface library, which currently provides AD computations of first and second derivatives for expressions in the AMPL modeling language (see http://www.ampl.com). Details behind this presentation are in the manuscript available at http://www.sandia.gov/~dmgay/bounds10.pdf.

6. <u>Eric Larour</u> (Jet Propulsion Laboratory, Pasadena CA, USA),
Jean Utke (University of Chicago /Argonne National Laboratory, USA),
Sri Hari Krishna Narayanan (Argonne National Laboratory, USA), Heather Cole-Mullen (University of Chicago, USA):
Automatic differentiation of the Ice Sheet System Model (ISSM) prognostic equations

Sea level rise is an issue of tremendous societal importance, which is presently poorly know or represented in Earth System models. In order to correctly assess what future sea level will be in a changing climate, the contribution of polar ice sheets must be thoroughly evaluated, and projected in time. The Ice Sheet System Model (ISSM) was built with this goal in mind. It was developed at the Jet Propulsion Laboratory, in collaboration with the University of California Irvine. It is capable of modeling continental ice sheets at high resolution, using a multi-model, multi-scale and multi-physics approach.
However, in order to spin-up transient ice flow models in a reliable way, a massive amount of satellite datasets must be assimilated into ISSM. This is usually done using inverse methods and adjoint-based optimization. Unfortunately, ISSM is coded entirely in C++, which presents challenges for the application of automatic differentiation compilers.
Here, we present the results of an initial effort to apply the ADIC2 compiler from Argonne National Laboratory to ISSM. We restrict our analysis to the prognostic equations of ice flow, which present a reduced set of model capabilities which we will build upon and extend in the near future. We demonstrate the applicability of ADIC2 to ISSM, and the advantages to be gained from automatic differentiation in spinning up ice flow models, as well as the implications for projections of future sea level rise.
This work was performed at the California Institute of Technology's Jet Propulsion Laboratory under a contract with the National Aeronautics and Space Administration's Cryosphere Science Program.

7. <u>Eric Phipps</u>, Roger Pawlowski, Andy Salinger (Sandia National Laboratories, Albuquerque NM, USA);
Roger Ghanem, Ramakrishna Tipireddy (University of Southern California, Los Angeles CA, USA) :
Embedded Stochastic Galerkin Propagation and Solver Methods via Template-based Generic Programming

Predictive computational simulation promises to help solve many important problems through- out science and engineering. A

critical task for predictive simulation is the quantification of uncertainties in simulation results arising from error or uncertainty in the input data. Thus there has been significant interest in developing computational methodologies for accurately and efficiently propagating representations of uncertainty in simulation input data to simulation output quantities of interest. While there are many approaches for representing uncertainty in simulations, in this work we are concerned with probabilistic representations of uncertain input data through random variables with prescribed probability distributions. Within this context, approaches fall into two categories: non-intrusive methods that sample the simulation code at various realizations of the input data, and intrusive or embedded methods that require more information from the simulation beyond quantity of interest realizations. In the latter case, a well-known class of such methods is intrusive stochastic Galerkin methods, which formulate new sets of equations governing the projection of simulation variables in spaces spanned by polynomials of the input random variables. The study and use of stochastic Galerkin methods in large-scale computational environments is hampered by their inherent intrusiveness, a given simulation code must be extensively reengineered to formulate and solve the stochastic Galerkin equations. In this work, we present a general approach based on template-based generic programming for automating much of the process of incorporating stochastic Galerkin methods in large-scale simulation environments. Template-based generic programming is a generalization of the ideas underlying automatic differentiation whereby computer code for a given calculation is transformed into code to compute additional derived information. Our approach leverages the template features of the C++ language by replacing the fundamental floating-point scalar type with carefully constructed classes that overload all of the floating point operations within the calculation. By combining this technique with traditional automatic differentiation, derivatives of stochastic Galerkin coefficients can be easily obtained. In this poster we describe a set of software tools called Stokhos, which is a package within the Trilinos solver framework. In conjunction with Sacado, a Trilinos package for automatic differentiation of C++ codes, Stokhos provides a toolkit for incorporating stochastic Galerkin methods in large-scale computational environments. Leveraging the template-based generic programming approach, we contrast several approaches for propagating stochastic Galerkin coefficients in the manner described above. Additionally, we describe several tools in Stokhos for assembling the stochastic Galerkin equations in large-scale distributed memory environments, interfaces that connect these equations to the numerous linear, nonlinear, transient, and optimization solvers within Trilinos, and customized linear solver methods leveraging the special structure of stochastic Galerkin methods. Finally we demonstrate how these techniques have been incorporated into several PDE simulation tools at Sandia, enabling ground-breaking research into large-scale embedded uncertainty quantification methods.

8. Ibraheem Alolyan (Mathematics Department, College of Science, King Saud University, Riyadh, Saudi-Arabia):
Exclusion Algorithm for Finding All Zeros of Vector Functions

Computing a zero of a continuous function is an old and extensively researched problem in numerical computation. In this paper we present an efficient subdivision algorithm for finding all real roots of a function in multiple variables. This algorithm is based on a simple computationally verifiable necessity test for existence of a root in any compact set. Both theoretical analysis and numerical simulations demonstrate that the algorithm is very efficient and reliable. Convergence is shown and numerical examples are presented.

9. Nikolai Strogies, Andreas Griewank, Daniel Josh (Humboldt University Berlin, Germany):
Optimal Sparse Matrix Compression via Lagrange-Newton Seeding

A significant improvement of matrix compression techniques for computing the Jacobian matrix $J(x)$ of a nonlinear vector function $F(x)$ is presented. To that end, the common Graph Coloring approach [Curtis-Powell-Reid] is combined with the method of defining a dense seed matrix S of minimal width such that the product matrix $J(x)S$ allows the reconstruction of all nonzero elements of $J(x)$ for a given sparsity pattern [Newsam-Ramsdell]. The Seed matrix is constructed from a generalized Vandermonde form via Lagrangian polynomials. The resulting compressed Jacobian $J(x)S$ can be approximated by difference quotients or evaluated with working accuracy by the forward mode of Automatic Differentiation. In either case the computational effort for computing the compressed Jacobian will be proportional to the evaluation cost of $J(x)$ times p, the number of columns in the matrix S. The construction of the compression scheme provides the property of p being equal to ϱ, the maximal number of nonzero elements in each row of the Jacobian. If this number coincides with the chromatic number χ of the column inter- section graph introduced by Coleman-More, the method reduces to the classical grouping scheme of Curtis-Powell-Reid. Otherwise all linear solving required is equivalent to polynomial interpolation on systems of order $(\chi - \varrho)$. Apart from the promising results concerning computational costs the question of numerical stability of the evaluation and reconstruction process is the main concern in the design and experimental tuning of the presented method.

10. James D. Turner, Ahmad H. Bani Younes, John L. Junkins (Texas A&M University, College Station TX, USA):
Generalized Least Squares Optimization for Nonlinear Observation Models

The existence of a nonlinear observation model creates the opportunity for investigating higher- order optimization algorithms. Classically a quadratic performance index is defined that is Taylor expanded to first order to create the necessary conditions required for differentially correcting an initial guess for the unknown system model parameters. This assumption implicitly assumes that second- and higher-order derivatives for the observation model are small. These assumptions are relaxed and generalizations are developed for extending the Least Squares Optimization to explicitly account for second- and higher-order derivatives. The nonlinear

necessary conditions are developed by introducing an artificial ordering parameter. The ordering parameter is used two ways: (1) the parameter correction, and (2) the observation model are Taylor expanded as a function of the unknown model parameters. Necessary conditions are collected as a function of the ordering parameter and sequentially solved to produce the desired correction vector solution. Automatic differentiation techniques are used to compute the higher-order observation model sensitivity partial derivatives. Classical weighted least-squares algorithms assume that a performance index if defined that seeks to minimize the errors arising between physical observations of a system and the assumed mathematical model for the system. The performance index is assumed to be a quadratic function of the observed minus the predicted system behavior. Given an initial guess for the unknown system parameters, a correction vector is developed by Taylor expanding the performance index to first-order in the parameter correction vector. The resulting necessary conditions are manipulated to provide the classical Least Squares algorithm. This paper extends this approach by introducing high-order expansion for both the math model and its gradient, yielding a higher-order necessary condition for defining the desired least squares solution. The first- through fourth-order partial derivative models for the assumed math model are obtained by using the recently developed Object-Oriented Coordinate Embedding (OCEA) method for automatic differentiation. The solution for the generalized least squares necessary condition is obtained by re-formulating the problem in a space of higher dimension, by introducing an artificial independent variable. The introduction of an artificial independent variable allows the current guess for the optimal parameter vector to be analytically continued in the higher-dimension space. The solution for the original problem is obtained by selecting special values for the artificial independent variable. The final solution is described by a series a reversion of series expansion.

11. Ahmad H. Bani Younes, James D. Turner (Texas A&M University, College Station TX, USA): Numerical Integration of Constrained Multi-Body Dynamical Systems Using 5th Order Exact Analytic Continuation Algorithm

Many numerical integration methods have been developed for predicting the evolution of the response of dynamical systems. An exact fifth-order analytic continuation method is presented for integrating constrained multi-body vector-valued systems of equations, where the Jacobi form of the Routh-Voss equations of motion simultaneously generates the acceleration and Lagrange multiplier solution. The constraint drift problem is addressed by introducing an analytic continuation method that rigorously enforces the kinematic constraints through five time derivatives. The proposed approach is expected to be particularly useful for stiff dynamical systems, as well as systems where implicit integration formulations are introduced. Numerical examples are presented that demonstrate the effectiveness of the proposed methodology.

12. Mihai Alexe, Adrian Sandu (Virginia Tech, Blacksburg VA, USA): A fully discrete framework for the adaptive solution of inverse problems

We investigate and contrast the differences between the *discretize-then-differentiate* and *differentiate-then-discretize* approaches to the numerical solution of parameter estimation problems. The former approach is attractive in practice due to the use of automatic differentiation for the generation of the dual and optimality equations in the first-order KKT system. The latter strategy is more versatile, in that it allows one to formulate efficient mesh-independent algorithms over suitably chosen function spaces. However, it is significantly more difficult to implement, since automatic code generation is no longer an option. The starting point is a classical elliptic inverse problem. An *a priori* error analysis for the discrete optimality equation shows consistency and stability are not inherited automatically from the primal discretization. Similar to the concept of dual consistency, we introduce the concept of *optimality consistency*. However, the convergence properties can be restored through suitable consistent modifications of the target functional. Numerical tests confirm the theoretical convergence order for the optimal solution. We then derive *a posteriori* error estimates for the infinite dimensional optimal solution error, through a suitably chosen error functional. These estimates are constructed using second order derivative information for the target functional. For computational efficiency, the Hessian is replaced by a low order BFGS approximation. The efficiency of the error estimator is confirmed by a numerical experiment with multiple grid optimization.

13. Mu Wang, Assefaw Gebremedhin, Alex Pothen (Purdue University, West Lafayette IN, USA): Automatic Differentiation Algorithms for Hessian Computation

For the purposes of derivative computation via Automatic Differentiation (AD), the evaluation of a function can be modeled using a directed acyclic graph (DAG) where vertices represent variables and edges represent data dependencies. Such a DAG, known as the computational graph of the function, is sufficient to work with for AD algorithms for Jacobian computation, but not for Hessian computation, since the graph fails to capture nonlinear interactions between variables. Gower and Mello [Optimization Methods and Software, 2011] introduced a graph model for Hessian computation in which the computational graph of the function is augmented with additional edges representing nonlinear interactions. This graph model exploits the symmetry available in Hessian computation. Using the model, they developed an algorithm, called edge_pushing, that exploits available sparsity 'on the fly'. An alternative approach of exploiting sparsity is to first determine the sparsity structure of the Hessian, and then compute the Hessian via compression. We report preliminary experimental results that compare the two approaches, sparsity exploitation on the fly versus sparsity exploitation via compression, on Hessians with various sparsity structures.

**14. <u>Drew Wicke</u> (George Mason University, Fairfax VA, USA),
Sri Hari Krishna Narayanan, Paul D. Hovland (Argonne National Laboratory, USA):**
Activity Analysis for Operator Overloading Automatic Differentiation

Often, AD users are interested in computing the derivatives of a subset of the output variables in the model with respect to a subset of the input variables (known as dependent and independent variables respectively). Depending on the role a temporary variable plays in the model source code, it may need to be involved in the computation of derivatives as well. In operator overloading AD, such variables usually are identified manually using a compiler that detects type mismatch. The overheads associated with operator overloading are exacerbated by the overestimation of such variables. Activity analysis is a two-step method to accurately identify those temporary variables that need to be involved in derivative computation. First, all variables whose value is affected by independent variables (vary variables) are detected. Second, all variables that affect the value of a dependent variable (useful variables) are detected. Variables that are both useful and vary are considered to be active. We have developed a framework based on the activity analysis implementation within Open-Analysis to detect active variables within code that was used as input to Sacado - an operator overloading AD tool. The types of variables that were determined to be active were changed to be of Sacado Active Type. The parsing of input code and type change was performed using the ROSE compiler framework and the useOA-ROSE interface. Preliminary experiments show that our tool is able to accurately distinguish those variables that must be active from those that should not.

15. <u>John Pryce</u> (Cardiff University, UK), Ned Nedialkov (McMaster University, Hamilton ON, Canada):
DAESA: a Matlab Tool for Structural Analysis of DAEs

DAESA, Differential-Algebraic Equations Structural Analyzer, is a Matlab tool for structural analysis of differential-algebraic equations (DAEs). It allows convenient translation of a DAE system into Matlab and provides easy-to-use functions to extract and to display structural information. DAESA can analyze systems that are fully nonlinear, high-index, and of any order. DAESA determines structural index, number of degrees of freedom, constraints, variables to be initialized, and suggests a solution scheme. The structure of a DAE can be readily visualized by this tool. DAESA can also construct a block-triangular form of the DAE, which can be exploited to solve it efficiently in a block-wise manner. The poster shows DAESA's capabilities by example.

16. <u>Michael Patterson</u>, Matthew Weinstein, Anil Rao (University of Florida, Gainesville FL, USA):
An Object-Oriented Method for Computing Derivatives of Mathematical Functions in MATLAB

An object-oriented method is presented that generates derivatives of functions defined by MATLAB computer codes. The method uses operator overloading together with forward mode automatic differentiation and produces a new MATLAB code that computes the derivatives of the outputs of the original function with respect to the variables of differentiation. The method can be used recursively to generate derivatives of any order that are desired and has the feature that the derivatives of a function are generated simply by evaluating the function on an instance of the class. The method is described briefly and is demonstrated on an example.

Last modified: July 20, 2012, AD 2012 Organizing Committee

Please check also the AD2012 conference web page at
http://www.autodiff.org/ad12