Invited Talks at AD 2012

Luca Capriotti (Credit Suisse, New York City, USA): Monday July 23rd, 1:30 pm
Algorithmic Differentiation: extending the frontiers of computational finance and risk management

Algorithmic differentiation has recently taken by storm the computational finance community, especially in the context of risk management. This is an area critically dependent on the ability to compute very quickly the sensitivities of the value of portfolios of securities with respect to a large number of market factors: a task traditionally tackled with a tremendous amount of computer power, at a high operational cost.   In this talk, I will introduce some of the keys applications of AD in finance and show how its tremendous computational effectiveness is allowing practitioners to significantly push the boundaries of what is practically achievable, also opening the way to more reliable risk management practices.
References:
- Luca Capriotti and Mike Giles, Algorithmic Differentiation: Adjoint Greeks Made Easy, preprint (2012).
- Luca Capriotti, Jacky Lee and Matthew Peacock Real Time Counterparty Credit Risk Management in Monte Carlo, Risk Magazine, May (2011).
- Luca Capriotti, Fast Greeks by Algorithmic Differentiation, Journal of Computational Finance 14, 3 (2011).
- Luca Capriotti and Mike Giles, Fast Correlation Greeks by Adjoint Algorithmic Differentiation, Risk Magazine, April (2010).

Markus Püschel (ETH Zurich, Switzerland): Tuesday July 24th, 8:30 am
Program Synthesis for Performance

Extracting optimal performance from modern computing platforms is becoming increasingly difficult. The effect is particularly noticeable in computations that are of mathematical nature such as those needed in media processing, communication, control, data analytics, or scientific simulations. In these domains, straightforward implementations often underperform by one or even two orders of magnitude. The reason is in optimizations that are known to be difficult and often impossible for compilers: parallelization, vectorization, and locality optimizations.

On the other hand, many mathematical applications spend most of their runtime in well-defined mathematical kernels such as matrix computations, transforms, interpolation, coding, and others. Since these are likely to be needed for decades to come, it makes sense to build program generation systems for their automatic production.

With Spiral we have built such a system for the domain of linear transforms. The key principles underlying Spiral are domain specific mathematical languages to represent algorithms, rewriting systems for different forms of parallelization and locality optimizations, and the use of machine learning to adapt code at installation time. Spiral-generated code has proven to be often faster than any human-written code and has been deployed in industry.

More information: www.spiral.net

Short Bio:
Markus Püschel is a Professor of Computer Science at ETH Zurich, Switzerland since 2010. Before, he was a Professor of Electrical and Computer Engineering at Carnegie Mellon University, where he still has an adjunct status. He received his Diploma (M.Sc.) in Mathematics and his Doctorate (Ph.D.) in Computer Science, in 1995 and 1998, respectively, both from the University of Karlsruhe, Germany. At Carnegie Mellon he was a recipient of the Outstanding Research Award of the College of Engineering and the Eta Kappa Nu Award for Outstanding Teaching. He also holds the title of Privatdozent at the University of Technology, Vienna, Austria. In 2009 he cofounded SpiralGen Inc. The work on Spiral has received several recognitions including being featured as NSF Discovery.

More information: http://www.inf.ethz.ch/~markusp

Andreas Griewank (Humboldt University Berlin, Germany): Tuesday July 24th, 1:30 pm
On Stable Piecewise Linearization and Generalized Algorithmic Differentiation

For several decades AD people and tools have uneasily stepped around the issue of nondifferentiabilities. Simultaneously, the nonsmooth analysis folks have build up a formidable array of generalized derivative concepts, and a magnificent theory in finite and infinite dimensional spaces to boot. The trouble is that from our low level 'every function is defined by a finite evaluation procedure' point of view, all that stuff never really happens. By that I mean that, once we are out in floating point country, all the weird and wonderful concepts reduce at almost all evaluation points x to plain vanilla gradients, Jacobians and Hessians. So why bother even implementing the first algorithm guaranteed to calculate generalized gradients and Jacobians (even conically active ones) presented by Khan and Barton at this conference? Of course, I claim to have the answer for exactly the same class of piecewise smooth functions Khan and Barton consider, namely compositions of smooth elementals plus abs(), min(), and max(). Piecewise linearization is a concept that is (in contrast to generalized differentiation) stable with respect to the base point x and it does yield information about nearby kinks, not just the one x sits on directly, which is very unlikely indeed as I keep stressing. Implementation is mostly a minute extension of standard AD tools. We give theorems that establish how unconstrained optimization, nonlinear equation solving, and numerical ODE integration can be based on piecewise linearizations of piecewise smooth objectives, vector functions and right hand sides.

Mary Hall (University of Utah, Salt Lake City UT, USA): Wednesday July 25th, 8:30 am
Autotuning Compiler and Language Technology and its Role in Exascale Systems

Predictions for exascale architectures include a number of changes from current supercomputers that will dramatically impact programmability and further increase the challenges faced by high-end developers. With heterogeneous processors, dynamic billion-way parallelism, reduced storage per flop, deeper and configurable memory hierarchies, new memory technologies, and reliability and power concerns, the costs of software development will become unsustainable using current approaches. In this talk, we will explore the limitations of current approaches to high-end software development, and how exascale architecture features will exacerbate these limitations. We will argue that the time is right for a shift to new software technology that aids application developers in managing the almost unbounded complexity of mapping software to exascale architectures. As we rethink how to program exascale architectures, we can develop an approach that addresses all of the productivity, performance, power and reliability concerns. We will consider proposed solutions from the community and how we might successfully deploy them. We will discuss one promising contemporary technology for mapping software to current complex hardware: autotuning, which employs empirical techniques to evaluate a set of alternative mappings of computation kernels to an architecture and select the mapping that obtains the best performance. When exascale systems demand optimizations with performance, energy and reliability goals, autotuning will become even more important, if the costs of searching a space of transformations can be managed.

Bert Speelpenning (MathPartners Inc., Seattle WA, USA): Wednesday July 25th, 11:00 am
An Idea Whose Time Had Come

Sometimes an idea is developed independently by different people at roughly the same time. This seems to have happened with what has become known as the reverse method. This talk offers personal observations and reflections on the birth of this idea, and the environment in which it came to fruition. The talk will conclude with a look into environments that are suitable for learning, and some implications for education, specifically math education.

Don Estep (Colorado State University, Fort Collins CO, USA): Thursday July 26th, 8:30 am
The use of adjoints for error estimation and uncertainty quantification

In this talk, we survey some uses of adjoint problems for aspects of uncertainty quantification for differential equations. We will discuss a posteriori error estimation, forward propagation of stochastic variation, and the formulation and solution of stochastic inverse problems for parameter determination.

Barbara Kaltenbacher (University of Klagenfurth, Austria): Thursday July 26th, 1:30 pm
Inverse Problems - Applications and Solution Strategies

The purpose of this talk is to give an overview on recent trends in regularization methods, with an emphasis on applications in the context of partial differential equations. After an introduction with some illustrating application examples we will highlight the special features and challenges of such inverse problems, especially their instability. The main part of this talk consists of a presentation of stable solution strategies - i.e., regularization methods - including recent trends such as regularization in Banach spaces (e.g., Lp). We conclude with some remarks on the resulting requirements on automatic differentiation in this context.

Lorenz Biegler (Carnegie Mellon University, Pittsburgh PA, USA): Friday July 27th, 8:30 am
Optimization of Pressure Swing Adsorption: A Case Study for Automatic Differentiation

Pressure Swing Adsorption (PSA) is an efficient method for gas separation, and is a potential candidate for carbon dioxide (CO_2) capture from power plants. However, few PSA cycles have been designed for this purpose and the optimization of PSA systems for CO_2 capture remains a very challenging task. In this study, we present a systematic optimization-based formulation for the synthesis and design of novel PSA cycles for CO_2 capture, which can simultaneously produce hydrogen (H_2) and CO_2 at a high purity. Here, we apply a superstructure-based approach to simultaneously determine optimal cycle configurations and design parameters for PSA units. This approach integrates ADOL-C, a powerful AD tool, efficient ODE solvers for the state and adjoint equations of the PSA model and state of the art NLP solvers. Three optimization models are proposed and two PSA case studies are considered. The first case study considers a binary separation of H_2 and CO_2 at high purity, where specific energy is minimized, while the second considers a larger five component separation. (paper co-authored with A. W. Dowling and S. R. R. Vetukuri)

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