Title: Progress in Higher Order Automatic Differentiation

Subtitle: AXIS OF EVAL! AUTOMATIC DIFFERENTIATION mates with LAMBDA CALCULUS birthing MONSTER COMPILER faster than FORTRAN

Professor Barak Pearlmutter, Maynooth University

Abstract:

The technique known in the machine learning community as "back-propagation" is a special case of "reverse-mode accumulation automatic differentiation", or "reverse AD". We will explore forward and reverse AD using novel formulations that make contact with differential geometry and the lambda calculus. In this context, the AD operators naturally generalize to a much broader range of computer programs, including programs containing iterate-to-fixed-point loops; invoking or embodying higher-order functions; invoking optimizers; or even themselves invoking AD operators. Algorithms including fast exact Hessian-vector multiplication, Pineda/Almeida fixed point back-propagation, and a wide variety of other techniques can be defined and implemented as one-liners. These methods allow very complicated systems, like bi-level optimization architectures, to be built and optimized using gradient methods. We are in the process of formalizing this system using the tools of Programming Language Theory, and a research prototype implementation has been constructed which exhibits startlingly good (faster-than-FORTRAN) numeric performance.