**475b Automatic Differentiation for Molecular Simulation**

*Derya B. Ozyurt and Paul I. Barton*

Any molecular mechanics (MM) and dynamics (MD) calculation to simulate a molecular system requires derivatives, specifically of the potential energy function. Although there have been efforts to improve the formulation of the analytical derivatives for MM and MD applications, accurate, efficient and automatic evaluation of the first and second order derivatives for these applications can be invaluable in several ways. For instance, an efficient Hessian-vector product calculation procedure coupled with a truncated-Newton type optimization algorithm can result in an effective structural refining of molecular systems, especially proteins. On the other hand, second order derivatives can improve the accuracy of numerical methods, such as symplectic integration for MD. Moreover, sensitivity analysis of a molecular simulation with respect to model parameters can also be performed using first order derivatives.

Automatic Differentiation (AD), different from ``symbolic'' and ``numerical'' differentiation methods, exploits the fact that a computer program evaluates any function as a sequence of elementary arithmetic operations by applying the chain rule repeatedly to these operations to obtain derivatives of arbitrary order automatically and without truncation error. Depending on how the chain rule is used to accumulate the derivatives through the computation, forward and reverse mode approaches for AD have been developed. The forward mode propagates derivatives of intermediate variables with respect to the independent variables and can be inefficient when there are many independent variables. In contrast, the reverse (or adjoint) mode accumulates derivatives by reversing the sequence of elementary arithmetic operations that evaluate the function. In this way, the gradient of a function involving many independent variables can be computed for a small multiple of the cost of a function evaluation, regardless of the number of independent variables. It should be noted that this reversal of the computation is normally implemented by automatic code transformation or operator overloading, and is not mimicked by careful hand coding of derivatives. Moreover, for a single function in terms of many independent variables, the correct combination of forward and reverse modes yields a Hessian-vector product for a small, constant multiple cost of the original function evaluation.

We propose an AD driven methodology for geometry optimization as an initial exercise to showcase the computational advantages of evaluating the gradients and Hessian-vector products using AD for molecular simulations in general. In particular, an existing implementation of several popular force fields will be differentiated using AD. The resulting computer programs evaluating gradients and Hessian-vector products will be utilized in a truncated-Newton method with an automatic preconditioner to obtain the minimum energy configuration of the example molecular systems. Effectively, by replacing the derivative evaluations based on applying the straightforward chain rule term by term (forward mode) with reverse mode AD, and by calculating Hessian-vector products directly instead of building a (sparse) Hessian matrix, the computational cost of MM can be reduced. To this end, computational comparisons with existing methods will be reported. Our investigations will also include possible improvements to MD especially for time stepping and parametric sensitivity analysis. Consecutively, we will present a methodology directed towards effective utilization of the accurate and efficiently generated derivative information from AD to study multi-scale modeling and simulation incorporating MM/MD and eventually quantum mechanical considerations.