Research Interest

Wei Wang

December 6, 2003

My research interest is numerical methods for computational biology with particular emphasis on *fast force-field evaluation and numerical integration methods* for biomolecular simulation. Biomolecular simulation provides insights into the biological structure and function at the microscopic level. It plays an increasingly important role in fundamental research as well as in practical applications such as drug design and medical research, partly due to the rapid advance in numerical algorithms and computer hardware, partly due to high experimental costs in some cases. The complexity of biological systems require larger and larger system sizes as well as longer and longer simulation times. Recent progress in numerical algorithms and parallel computing has already made possible the simulation of a two-million atom system in a nanosecond time scale (1ns = 10^{-9} second) in a recent study on human ribosome.

I am currently completing my Ph.D. research on *developing fast algorithms for computing polarizable forces* under the guidance of my advisor, Professor Robert D. Skeel in the Computer Science Department at the University of Illinois at Urbana-Champaign. A point dipole model is a natural extension of the point charge model to describe more accurately the electrostatic interactions between atoms in biomolecular systems. Polarization is believed to play a crucial role in some biological processes, which cannot be explained in the framework of point charge models. Compared with other types of interactions, polarization is non-additive, meaning that a self-consistent solution has to be found through solving a large set of linear equations. This makes the computation very expensive, and hinders the development of the polarization models. For example, the current reliable Amber implementation of the point-dipole model is 6 times more expensive than that without dipoles. The goal of my Ph.D. research is to develop fast polarization solvers so that polarizable force computation becomes a routine part of biomolecular simulation. My current algorithm has already reduced this cost by a factor of 2. Ongoing investigations into better preconditioners and predictors will help to reduce the cost even more. The practical culmination of my work will be the implementation of the fast point-dipole solver in the parallel molecular dynamics program *NAMD*, a Gordon Bell Award winner at the 2002 Supercomputing Conference. NAMD is currently the most scalable molecular dynamics software for biomolecular simulations. It is developed by the Beckman Institute Theoretical and Computational Biophysics Group, of which I am a member.

My M.S. thesis research was on *numerical integration methods for Langevin dynamics*. The time and size scale of biomolecular diffusion is generally too large to simulate by Newtonian dynamics. The challenge is partly relieved by treating the solvent (e.g., water) implicitly and integrating the Langevin equation to obtain "typical" trajectories to sample/traverse phase space. Theory for numerical integrators for the Langevin equation, a *stochastic ordinary differential equation*, has not been as well developed as that for Newtonian dynamics, for which symplectic integrators are the well-established methods to use due to their enhanced long time stability. Backward error analysis

has proved very fruitful for numerical integration methods for Newtonian dynamics. My work is the first investigation applying backward error analysis to stochastic differential equations. This work reveals that the popular Brooks-Brünger-Karplus method has weak convergence only of order 1, while some other methods have order 2 and are more accurate.

I was initially trained as a *physicist* and did some research in that area. I used the *Density Matrix Renormalization Group* (DMRG) method to compute the impurity energy levels for a quantum spin-1 chain. The DMRG method is an extension of the renormalization group technique widely embraced for the study of phase transitions in physics, and it has been generalized to many other areas. My result settled a disagreement on the nature of the impurity energy levels. I also implemented the *hyperdynamics* method to speed up the time scale of molecular dynamics simulation for studying the diffusion of ad-atoms on silicon (1 0 0) surface. Hyperdynamics is a novel method trading the uninteresting local-minimum oscillations for the interesting barrier hopping events while keeping the time scale correct in the statistical sense. Ad-atom diffusion was greatly enhanced by this method.

There are several natural extensions of my Ph.D. research. Other models for polarizable force fields, such as Drude models and fluctuating charge models, may assume importance, and their computation should be explored. It is also interesting to investigate the dynamical variable methods, such as the well-known Car-Parrinello method for the polarization computation. These methods treat the dipoles as dynamical variables and do the optimization on the fly with much smaller cost. It is very possible that the preconditioning, prediction, and acceleration techniques developed in my research are generalizable to other problems in computational biology. Other interesting directions in biomolecular simulation include the Poisson-Boltzmann equation in molecular simulations, hybrid quantum mechanics/molecular mechanics, time-stepping, sampling techniques, finite-element methods for elastic models of biomolecules, and multi-scale techniques. Beyond the scope of biomolecular simulations, I am also happy to collaborate with researcher in other fields, such as chemistry and material science, to do multidisciplinary research.

Teaching Philosophy

An important reason for me to pursue an academic career is that I enjoy the interaction between teacher and student. As a graduate student, I have taken two teaching enrichment classes and was in charge of lab sessions twice a week for two quarters. I also served as lab assistant in the summer school organized by our Theoretical and Computational Biophysics Group. The experience brought me much enjoyment and satisfaction in conveying knowledge and communicating with students.

A very important factor of effective teaching is to present the concepts in a way that fits students' backgrounds. *Illustration* is a good technique to use. For example, when I gave a class presentation on Langevin dynamics, which is unfamiliar to the audience, I provided, at the beginning, a few examples of Brownian motion to illustrate how the random and friction forces come into play. After that, the audience accepted the Langevin equation very easily without any complicated theoretical derivation and approximations.

A higher goal for the teacher is to help the students *achieve new perspectives through the use of the concepts being taught*. When I taught a physics lab a few years ago, I usually started with a ten-minute lecture to give an overview. I found it was particularly effective when I applied concepts to their daily lives. Questions, such as "what is the walking speed of a human?," "how fast can a human being run?," and "can you compare it with the speed of automobiles?," are well received and enlightening. Actually, presenting the knowledge in a different way is rewarding to the teacher too. New insights have been gained from my own lecturing experience.

Good presentation skills is key to effective classroom teaching. During the years, I have given various presentations in conferences, workshops, seminars, and group meetings. And my presentation skill has improved a lot through these experiences. Some general rules of thumb, such as "figures are more effective than words," "one slide, one idea," and "repeat key points" are well followed. What is more, I realize there is a difference between research-oriented presentations and instruction-oriented lecturing: the former emphasizes ideas; while in the latter, details must be well taken care of. Classroom lecturers should demonstrate clearly the process of thinking and derivation, especially for undergraduate teaching.

A good teacher should be aware of audience reaction and seek feedback. After a not-quite-successful presentation I gave a few years ago, I realized that if I lost the audience at the start, the rest of the presentations would be very ineffective. The way to avoid it is observe the audience and get feedback early. After the first few slides is time to solicit questions. If the reaction is not good, going back to earlier slides is a better choice than continuing blindly.

In summary, although my teaching experience is limited, I believe my enthusiasm and attitude would make me a good teacher.