Incorporating the GAMESS Program Into the GAML Software

Santiago Cortes[1], **Thomas Spangler**[2], Orlando Acevedo[3]

[1]Chemical and Biomolecular Engineering, Johns Hopkins University, [2]Computer Science, University of North Carolina Wilmington, [3]Department of Chemistry, University of Miami

Background & Significance

The Genetic Algorithm Machine Learning (GAML) software was created by Professor Orlando Acevedo and his graduate student Xiang Zhong to automate the creation of force field (FF) parameters for use in molecular dynamics (MD) or Monte Carlo (MC) simulations. These simulations can be used to predict precise macroscopic solvent properties, such as density, heat of vaporization, and free energy of hydration, for any desired specific molecule. GAML uses a quantum mechanical (QM) program called Gaussian in its code to compute partial atomic charges for use in FF parameterization. However, a major drawback is that while Gaussian is a powerful program, it is an expensive commercial software package. Our project goal was to integrate the code of an alternative open source QM software known as the General Atomic and Molecular Electronic Structure System (GAMESS) into GAML to calculate partial atomic charges. With the addition of GAMESS, GAML would become a completely open source program expanding accessibility to more users.

Objectives & Approach

In this work, new Python code was created and added to the GAML software to integrate the GAMESS software. This code first locates a Gaussian input file (.com) containing the user desired QM level of theory and the Cartesian coordinates that define the geometry of the solvent molecule and extract that data to create an input file for GAMESS (.inp). After submission and completion of the GAMESS calculation, the new code would then parse the output file and return atomic partial charges e.g., CHELPG charges, needed for the GAML code to proceed with the genetic algorithm-based FF parameterization of the solvent.

Outcomes & Future Directions

The project successfully integrated the new Python script into the GAML framework. Our work correctly located the specific information in the Gaussian input file, created a new GAMESS file, submitted the GAMESS calculation, and updated GAML with the newly predicted atomic partial charges. For future work, the GAML program needs to implement additional functionality for use in the validation of solvent properties. First, we need to implement radial distribution functions (RDF) verification, or the ability for the program to quantify the strength and structure of the intermolecular interactions present between solvent molecules in the bulk phase. Second, GAML needs to be expanded to parameterize the nonbonded Lennard-Jones (LJ) terms of a FF. The LJ terms play a major role in RDF predictions. An additional direction would be to create a web-server version of GAML that would expand accessibility to users with limited experience in Python, Linux, and command-line software.

Name: Santiago Cortez
School: Johns Hopkins University
Major: Chemical and Biomedical Engineering
Department: Department of Chemistry
Mentor: Dr. Orlando Acevedo

Name: Thomas Spangler
School: University of North Carolina Wilmington
Major: Computer Science
Department: Department of Chemistry
Mentor: Dr. Orlando Acevedo