ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

High Performance Computing for Science and Engineering II

Spring semester 2017

Dr. P. Hadjidoukas, Prof. C. Papadimitriou
ETH Zentrum, CLT E 13
CH-8092 Zürich

Project # 4

Issued: 15.05.2017

- In these scripts, we outline five computational engineering problems. Choose one and work either individually or in a group of two people. Communication is allowed to the extent that you do not copy the work of other groups.
- You are encouraged to contact one of the TAs in order to arrange a meeting to discuss your chosen project. These meetings are meant for clarifying and detailing the projects, give early feedback on your approach, *not* for correcting your code.
- In evaluating your work, we will consider your ability to analyze the problem and the hardware at your disposal, to appropriately apply the principles taught during the whole HPCSE class, and to report your reasoning and findings.
- The report (including text, code, figures) needs to be emailed before the day of the exam. If working in a group of two, each student has to write an *individual* report.

Global Optimization of Lennard-Jones Clusters

Numerical global optimization of real valued functions is an indispensable methodology for solving a multitude of problems in science and engineering. Many problems exhibit a number of local and/or global minimizers, expensive function evaluations or require real-time response.

Atomic Clusters using Pairwise Energy Potentials

Pairwise energy potentials are mathematical models used to calculate the total energy in a cluster of atoms. Minimizing this energy with respect to the coordinates of the atoms, corresponds to finding a stable conformation of the cluster. In the case of the pairwise potentials, the total energy of the system consisting of N_{atoms} atoms can be calculated as:

$$U_{tot} = \sum_{i=1}^{N} \sum_{j<i}^{N} E\left(r_{ij}\right) \tag{1}$$

Here the quantity $E\left(r_{ij}\right)$ is the pairwise interaction energy between i-th and j-th atom and

$$r_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i - z_j\right)^2}$$

is the distance between atoms i located in $(x_i, y_i, z_i)^\top$ and atom j located in $(x_j, y_j, z_j)^\top$. Different formulations for $E(r_{ij})$ lead to different potentials and hence different conformations in space. Researchers over the last years proposed different pairwise energies to describe interactions for various classes of atoms.

Lennard Jones Potential

The Lennard-Jones potential [1] is a simple model that approximates the interaction between a pair of neutral atoms or molecules. It is considered a relatively good and universal approximation and due to its simplicity is often used to describe the properties of gases.

Formation of Lennard-Jones clusters have been extensively used as global optimization benchmark problems. The potential energy of the cluster is given by:

$$E(r_{ij}) = 4\,\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right], \tag{2}$$

where r_{ij} is the distance between atoms i and j.

In this project we will use reduced units, i.e., $\sigma = \varepsilon = 1$. Moreover, all particles are within the range of [-1, 1], as described in [4, 2]. The list of global minima for this problem can be found at `http://doye.chem.ox.ac.uk/jon/structures/LJ.html`[1].

Global Optimization Algorithms

We consider two global optimization algorithms for solving the previously described problem:

- CMA-ES: Covariance matrix adaptation evolution strategy (CMA-ES) is a powerful population-based stochastic evolutionary algorithm for complex numerical optimization. CMA-ES user a multivariate normal distribution to model the population and offspring are generated by sampling from this distribution.
- DIRECT: The DIRECT algorithm [3] belongs to the class of direct search algorithms. The principle of the algorithm is described by its name, which is an acronym for DIviding RECTangles.

According to the DIRECT algorithm, the search space is divided into non-overlaping hyperrectangles and the middle (base) point of each hyperrectangle is evaluated. A hyperrectangle is characterized by the fitness of its base point and its size. The algorithm uses these characteristics to decide which of the existing hyperrectangles should be split at each iteration. As a starting point, we provide an open-source serial implementation of the algorithm in C, available from `http://www.dis.uniroma1.it/~lucidi/DFL/`[2].

[1] Reference implementation: `http://cse-lab.ethz.ch/images/teaching/HPCSEII_FS2017/LJ.zip`

[2] `http://cse-lab.ethz.ch/images/teaching/HPCSEII_FS2017/DIRECT.zip`

Problem setup

Your tasks in this project are the following:

- Parallelize the DIRECT optimization method on shared and distributed memory systems using the OpenMP and MPI programming models.
- Extend the available parallel implementation of the CMA-ES optimization algorithm on GPUs using the CUDA programming model. You can use an appropriately large number of samples to evaluate the performance of your implementation.
- Develop an optimized code that computes the total energy of Lennard-Jones clusters on the above platforms.
- Verify the correctness and evaluate the performance of your implementations by finding the optimal configurations for various number of clusters.

References

[1] https://en.wikipedia.org/wiki/Molecular_dynamics#Potentials_in_MD_simulations.

[2] W. Cai, H. Jiang, and X. Shao. Global optimization of lennard-jones clusters by a parallel fast annealing evolutionary algorithm. *Journal of Chemical Information and Computer Sciences*, 42(5):1099–1103, 2002.

[3] D. R. Jones, C. D. Perttunen, and B. E. B. E. Stuckman. Lipschitzian optimization without the lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181, 1993.

[4] D. J. Wales and J. P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *J. Phys. Chem*, 101:5111–5116, 1997.