CS9670B: Project proposal (Carter Wilson and Ryan Divigalpitiya)

What is the problem?

Molecular dynamics simulations allow one to model the motions of particles in a system and extract important thermodynamic and kinetic information. A key determinant of simulation accuracy is the quality of the force field used. A force field represents the QM potential energy surface using simple functional forms, in particular the Lennard-Jones approximation for particle-particle interactions. This potential is given by

$$U_{\mathrm{LJ}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$

where ε and σ are the interaction strength and size of two spherical particles at some distance r. This potential has been used to model a vast array of molecules and particles, including the noble gases. Optimization of ε and σ is a non-trivial task. We want to *optimize the LJ-potential using the structure factor of argon gas and cross-validate against diffusion data*. Subsequently, we will investigate whether transfer learning can be employed to tackle the same problem for neon gas.

One can conceive of the problem as an exploration of the 2D-parameter space $\mathbb{S}^2 = R_\varepsilon \times R_\sigma$, where the agent's goal is to find the cell $(\varepsilon_i, \sigma_j)$ that minimizes $G(\varepsilon_i, \sigma_j)$:

$$\mathrm{G}(\varepsilon,\sigma) = \sqrt{\frac{1}{n}\sum_{k}^{n} \mathrm{F}_k^{\mathrm{Experiment}} - \mathrm{F}_k^{\mathrm{Simulation}}(\varepsilon,\sigma)};$$

that is, the root mean squared deviation between the experimental and computed structure factor. For each step a penalty is accrued, the termination and the maximal reward is given for the optimal (ε, σ) pair, and 50 steps represent an episode.

What reinforcement learning techniques do you plan to experiment with?

We aim to consider two different RL-strategies. In the base case we have the 2D, (ε, σ) state-space, with corresponding rewards $r = [\mathrm{G}(\varepsilon,\sigma)]^{-1}$; actions (i.e., 1 or 2-steps in parameter space) move the agent within this parameter space. Given this is an MDP we will employ Q-learning and once converged a policy-transfer will be attempted to a different noble gas system, neon. A second setup will treat the state as an N-dimensional vector containing the numeric difference between the current structure factor and experiment, while the actions remain the same. Here an attempt at deep Q-learning will be attempted given the larger state space. In both cases performance will be compared to a standard genetic algorithm.

References

(1) Chatterjee, P. et al. *J. Chem. Theory Comput.* **2022**, *18*, 2388–2407.
(2) Rutkai, G. et al. *Mol. Phys* **2016**, *115*, 1104–1121.
(3) Mnih, V. et al. **2016**, DOI: 10.48550/ARXIV.1602.01783.
(4) Mnih, V. et al. *Nature* **2015**, *518*, 529–533.
(5) Mnih, V. et al. **2013**, DOI: 10.48550/ARXIV.1312.5602.
(6) Zhu, Z. et al. **2020**, DOI: 10.48550/ARXIV.2009.07888.