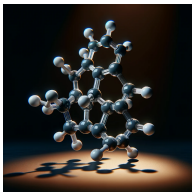


Predicting Atom Interactions with AI

Ludwig Schneider

Eric and Wendy Schmidt AI-Postdoctoral Fellow
Pritzker School of Molecular Engineering
University of Chicago

April 17, 2024



Molecular Interaction: Non-Bonded I

- ▶ Newtons Equation of Motion

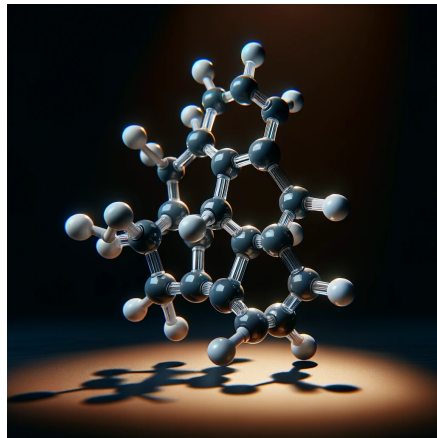
$$m_i \frac{d^2}{dt^2} \vec{r}_i = \vec{F}_i$$

- ▶ determine forces \vec{F}
- ▶ OPLS: functional form for the forcefield

$$V_{\text{nb}}(r_{ij}) = \epsilon_{ij} \left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}} - \frac{\sigma_{ij}^6}{r_{ij}^6} \right) + \frac{q_i q_j e^2}{4\pi\epsilon_0 r_{ij}}$$

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j} \quad \sigma_{ij} = \sqrt{\sigma_i \sigma_j}$$

Challenge: determine ϵ_i , σ_i , q_i , and m_i

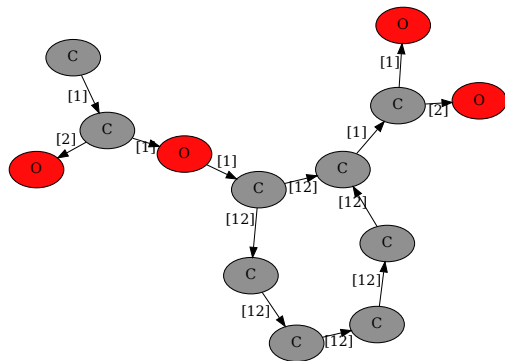


Molecular Interaction: Non-Bonded Interactions II

- ▶ determine ϵ_i , σ_i , q_i , and m_i
 - ▶ element (carbon, hydrogen)
 - ▶ molecular environment
- ▶ we know some values
- ▶ SMARTS based rules
- ▶ Quantum Mechanic simulations
- ▶ chemical space is vast, we want to interpolate

Your Challenge

Use ML to interpolate from known molecules!



Details of the Basic Challenge

- ▶ 2000 molecular graphs
- ▶ known input parameters
- ▶ supervised learning possible
 - 1 predict non-bonded parameters
 - 2 predict uncertainty in interpolation
- ▶ you are free to choose your model
- ▶ you are free to choose tech stack
- ▶ <https://github.com/badeaa3/UChicago-AI-in-Science-Hackathon>
- ▶ [molecular/intro/talk.pdf](#) this slide deck
- ▶ [molecular/README.md](#) starting point
- ▶ [molecular/explain_graph_data.py](#) explanation of data set
- ▶ [molecular/data.json](#) training data
- ▶ molecular graph as `networkx` graphs

How we evaluate you

- ▶ prepare a final presentation
- ▶ convince us your solution is best
- ▶ include model details and results
- ▶ XX minutes to present
- ▶ YY minutes for our questions
- ▶ morning last day: we give you additional data
- ▶ DO NOT TRAIN with this data
- ▶ `molecular/final_evaluation.py`
- ▶ report output in your presentation

What if you solve the challenge too fast?

Traditional

- ▶ predict bond class for atom
- ▶ look-up table for parameters
- ▶ `ffbonded.itp`

GNN approach

- ▶ predict parameters directly
- ▶ bonds: edge features
- ▶ angles & dihedrals ?

- ▶ bonds: 2 atoms: 2 parameters

$$V_b(r_{ij}) = k_a(r_{ij} - b_0)^2$$

- ▶ angles: 3 atoms: 2 parameters

$$V_a(\theta) = k_\theta(\theta - \theta_0)^2 \quad \theta = \cos^{-1}(\vec{r}_{ij} \cdot \vec{r}_{jk} / (||\vec{r}_{ij}|| ||\vec{r}_{jk}||))$$

- ▶ dihedral: 4 atoms: 8 parameters

$$V_d(\phi) = \sum_{i=1,2,3,4} V_i/2 [1 - (-1)^i \cos(i \cdot \phi - \phi_i)]$$

Happy Hacking!