Introduction:
Learning molecular energy functions

CS/CME/Biophys/BMI 371
Feb. 1, 2018
Ron Dror
A typical molecular mechanics force field

\[ U = \sum_{\text{bonds}} k_b \left( b - b_0 \right)^2 + \sum_{\text{angles}} k_\theta \left( \theta - \theta_0 \right)^2 + \sum_{\text{torsions}} \sum_{n} k_{\phi,n} \left[ 1 + \cos \left( n\phi - \phi_n \right) \right] \]

- **Bond lengths ("Stretch")**
- **Bond angles ("Bend")**
- **Torsional/dihedral angles**

\[ + \sum_{i} \sum_{j>i} \frac{q_i q_j}{r_{ij}} \]

- **Electrostatic**

\[ + \sum_{i} \sum_{j>i} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \]

- **Van der Waals**

**Bonded terms**

**Non-bonded terms**
Example: Bond length stretching

- A bonded pair of atoms is effectively connected by a spring with some preferred (natural) length. Stretching or compressing it requires energy.

\[ U(b) = k_b (b - b_0)^2 \]
Example: Electrostatics interaction

- Like charges repel. Opposite charges attract.
- Each atom carries some "partial charge" (may be a fraction of an elementary charge), which depends on which atoms it’s connected to.

\[ U(r) = \frac{q_i q_j}{r} \]

where \( q_i \) and \( q_j \) are partial charges on atoms \( i \) and \( j \).
Could we learn an energy function (force field)

- What if instead of writing the force field as a sum of terms each of which makes physical sense, we represent it as a large neural network?
  - We can then train that network on the results of many quantum chemistry computations
- Researchers have been working on this for over a decade, but it’s picked up steam in the last couple year
- One of Tuesday’s papers (Smith et al.) reports a substantial step in this direction
Two related ideas

• Another paper (Faber et al.) explores prediction of chemical properties of small molecules by machine learning
  – Learning is again based on quantum chemistry results, but there’s no force field involved

• A third paper (Park et al.) discusses improvement of the Rosetta all-atom force field by fitting to a wider variety of data types
Background material

• Introduction to energy functions (force fields) from CS/CME/BioE/Biophys/BMI 279:

• Discussion of the Rosetta force fields from CS/CME/BioE/Biophys/BMI 279: