

MDXを活用した物質・材料シミュレーションと情報共有

奥村雅彦

日本原子力研究開発機構
システム計算科学センター
シミュレーション技術開発室



物質・材料シミュレーション手法(ミクロ)

- 古典分子動力学法
- 第一原理計算(密度汎関数法)
- **機械学習分子動力学法**

良い点 / 悪い点

	計算対象	原子間相互作用	特徴
古典分子動力学	質点 (原子/イオン)	原子位置の解析関数	<ul style="list-style-type: none"> ・ 経験的パラメーターが必要 ・ 高速/低精度
第一原理分子動力学 (密度汎関数法)	密度(電子) 質点(原子核)	電子密度と原子核位置の関数	<ul style="list-style-type: none"> ・ 経験的パラメーター不要 ・ 低速/高精度
機械学習分子動力学	質点 (原子/イオン)	ニューラルネットワーク等	<ul style="list-style-type: none"> ・ 経験的パラメーター不要 ・ 高速/高精度

物質・材料シミュレーション手法

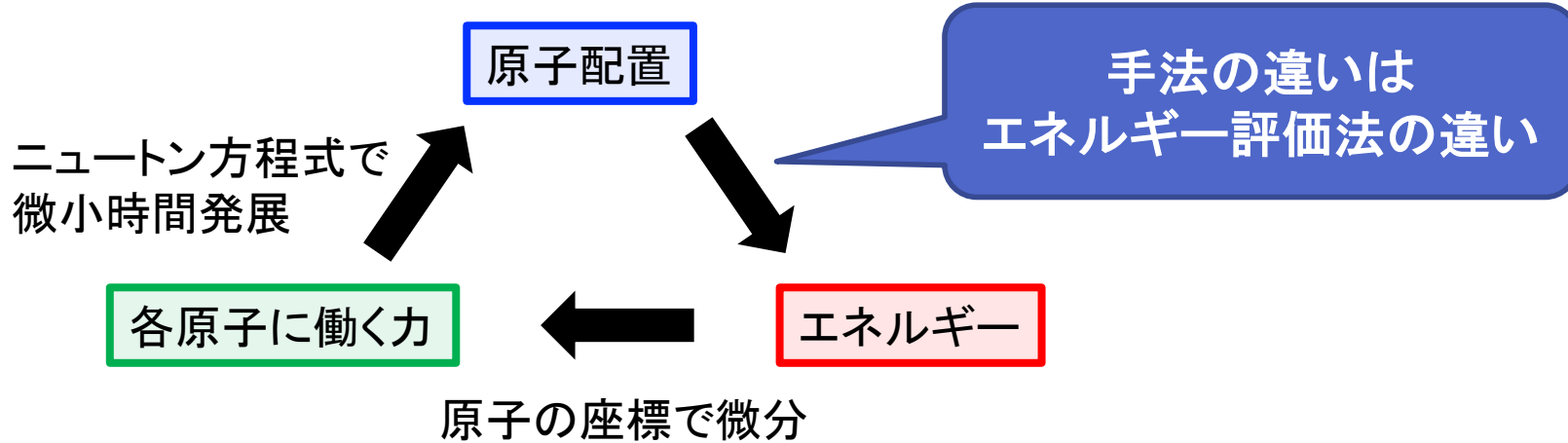
- 古典分子動力学法
- 第一原理計算(密度汎関数法)
- **機械学習分子動力学法**
 - J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007)

Behler-Parrinello 2007 年間引用数



分子動力学法とは

- 分子動力学法(Molecular Dynamics, MD)



- エネルギー評価法

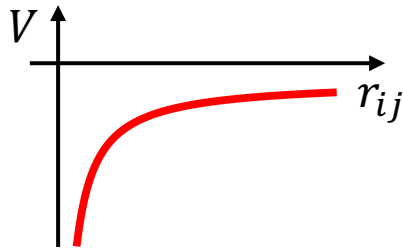
- 古典分子動力学法 (Classical MD)
- 第一原理分子動力学法 (*Ab initio* MD, First-principles MD)
- 機械学習分子動力学法 (Machine Learning MD) ← **NEW!**

古典MDにおけるエネルギー評価法

- 原子を質点、原子間力を簡単な解析関数で近似

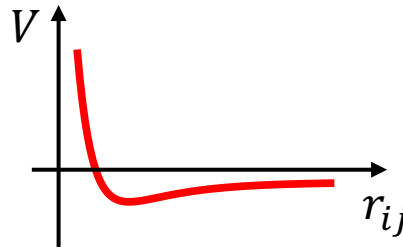
Coulombポテンシャル

$$V_C = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$



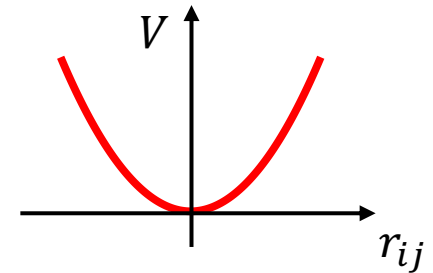
Lennard-Jonesポテンシャル

$$V_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right]$$



調和ポテンシャル

$$V_h = \frac{k}{2} (r_{ij} - r_0)^2$$



特徴

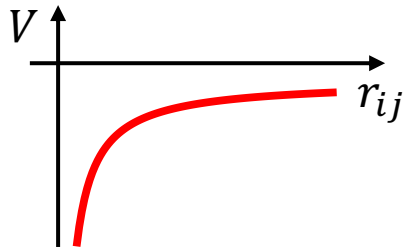
- いくつかの実験値や第一原理計算結果を再現するようにパラメーターを調整
- 調整には“職人技”が必要(最近は自動で作成する手法も登場)
- 計算は軽い**

古典MDにおけるエネルギー評価法

- 原子を質点、原子間力を簡単な解析関数で近似

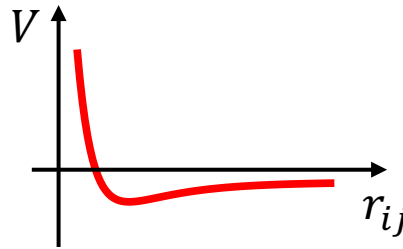
Coulombポテンシャル

$$V_C = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}$$



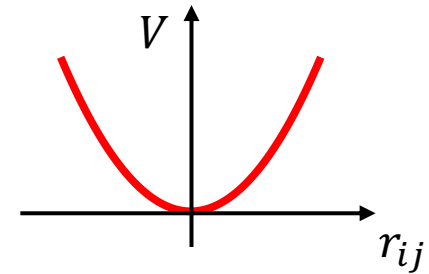
Lennard-Jonesポテンシャル

$$V_{LJ} = \epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right]$$



調和ポテンシャル

$$V_h = \frac{k}{2} (r_{ij} - r_0)^2$$



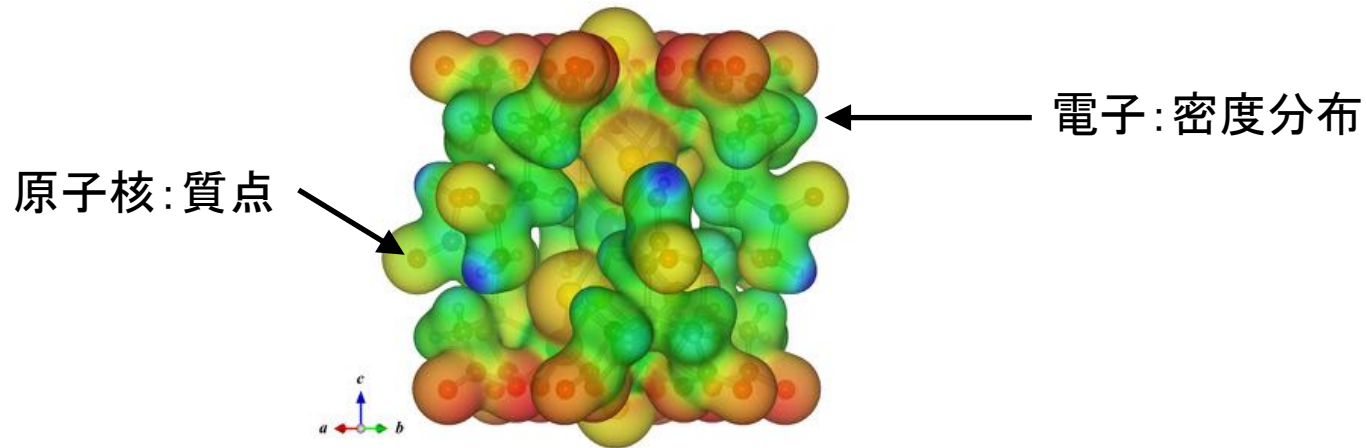
特徴

- いくつかの実験値や第一原理計算結果を再現するように **パラメーター** を調整
- 調整には“職人技”が必要(最近は自動で作成する手法も登場)
- 計算は軽い**

第一原理計算におけるエネルギー評価法

- 密度汎関数法

- 原子核は質点(古典力学)に近似するが、電子は量子力学を解く。



<http://jp-minerals.org/vesta/jp/features.html>

- 特徴

- 経験的なパラメーターは含まない
- “職人技”は必要ない
- 計算は重い

古典MDと第一原理MDの違い

ポテンシャルエネルギー曲面

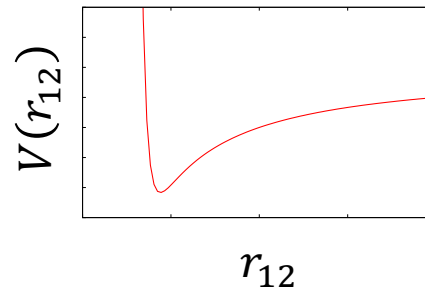
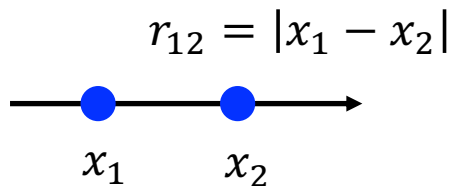
- 原子配置を引数とする関数: $V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$
 - 古典MD (2原子間距離のポテンシャルで近似): 引数1次元の関数

$$V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \cong \sum_{i < j} V(r_{ij})$$

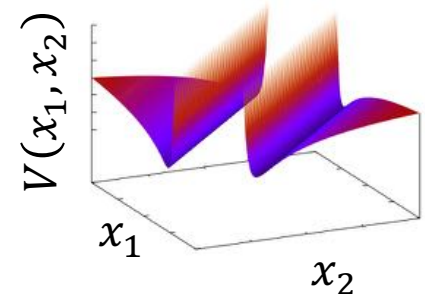
- 第一原理MD: 引数 $3N$ 次元の関数

$$V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$$

例: 1次元2原子系



古典MD

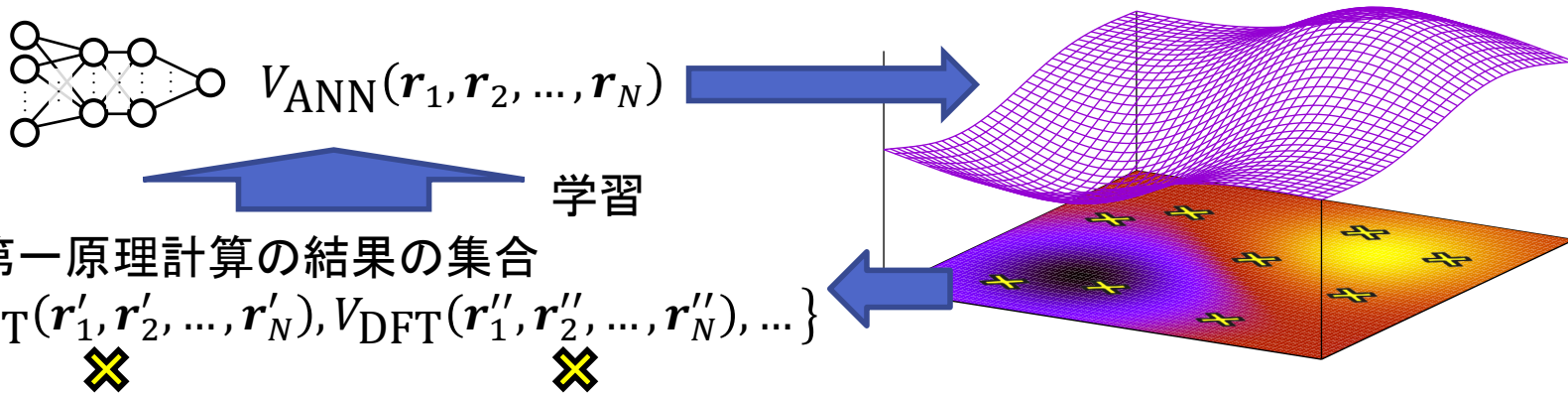


第一原理MD

機械学習MDのポテンシャルエネルギー曲面

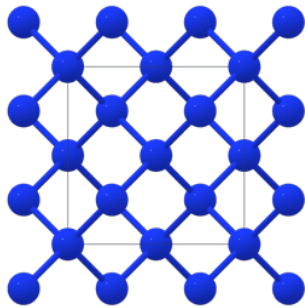
• 機械学習MDとは？

- 人工ニューラルネットワーク等で**第一原理計算のポテンシャルエネルギー曲面**を近似する手法
 - 人工ニューラルネットワーク: 任意の関数を近似可能な手法
 - **第一原理計算のポテンシャルエネルギー曲面**
 - 引数 $3N$ 次元の関数
 - 簡単な関数で書くことができない
- ➡ データセットを作成して人工ニューラルネットワークで学習する

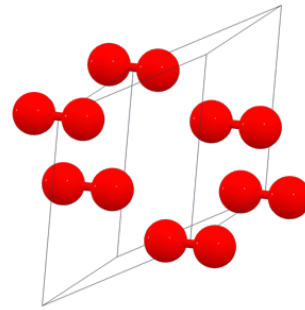


機械学習MDに必要なデータセット

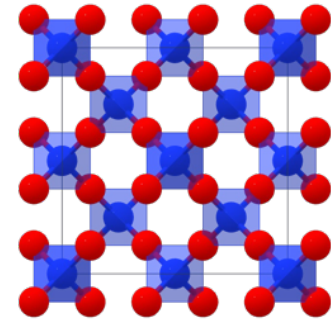
- 物質ごとにデータセットが必要
 - 例: Siのデータセット、Oのデータセットを持っていても、 SiO_2 のポテンシャルエネルギー曲面は構成できない



Si



O

 SiO_2

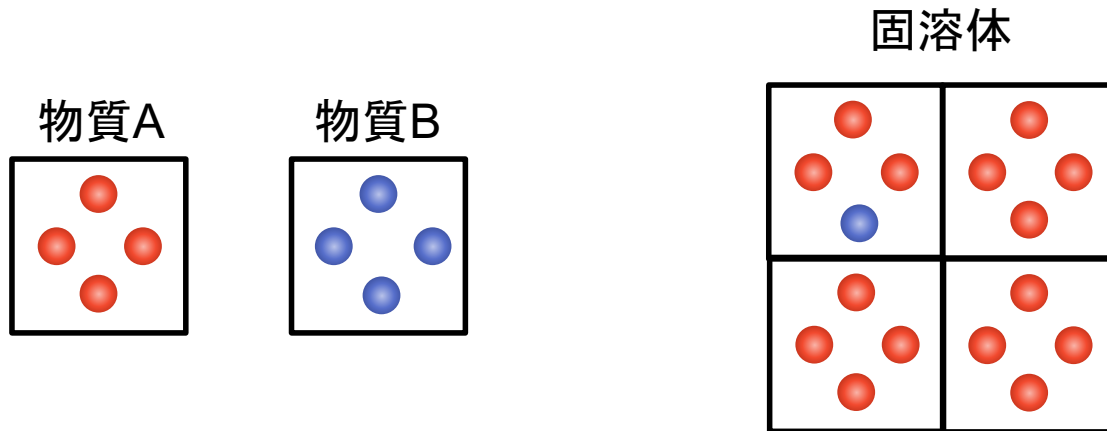
Siのデータセット: $\mathcal{V}_{\text{Si}} = \{V_{\text{Si}}(\mathbf{r}'_{\text{Si}1}, \mathbf{r}'_{\text{Si}2}, \dots, \mathbf{r}'_{\text{Si}N}), V_{\text{Si}}(\mathbf{r}''_{\text{Si}1}, \mathbf{r}''_{\text{Si}2}, \dots, \mathbf{r}''_{\text{Si}N}), \dots\}$

Oのデータセット: $\mathcal{V}_{\text{O}} = \{V_{\text{O}}(\mathbf{r}'_{\text{O}1}, \mathbf{r}'_{\text{O}2}, \dots, \mathbf{r}'_{\text{O}N}), V_{\text{O}}(\mathbf{r}''_{\text{O}1}, \mathbf{r}''_{\text{O}2}, \dots, \mathbf{r}''_{\text{O}N}), \dots\}$

SiO_2 のデータセット: $\mathcal{V}_{\text{SiO}_2} = \{V_{\text{SiO}_2}(\mathbf{r}'_{\text{Si}1}, \mathbf{r}'_{\text{O}1}, \mathbf{r}'_{\text{O}2}, \dots), \dots\} \neq \mathcal{V}_{\text{Si}} \cup \mathcal{V}_{\text{O}}$

機械学習MDにおけるデータ共有の有用性

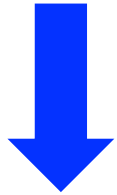
- 第一原理計算は計算コストが高い
 - 数万原子配置のデータセットが必要
 - 先行研究と同じ物質の機械学習MDをやるために膨大な第一原理計算を繰り返したくない
- 固溶体や界面の計算に有用



教師データの共有が有用

統合機械学習分子動力学システムの構築

- 2022年度JHPCN課題(jh220057)
- 目的: 誰でも機械学習MD計算ができるようにしたい
 - 機械学習MDのための教師データのデータベース



教師データの利用

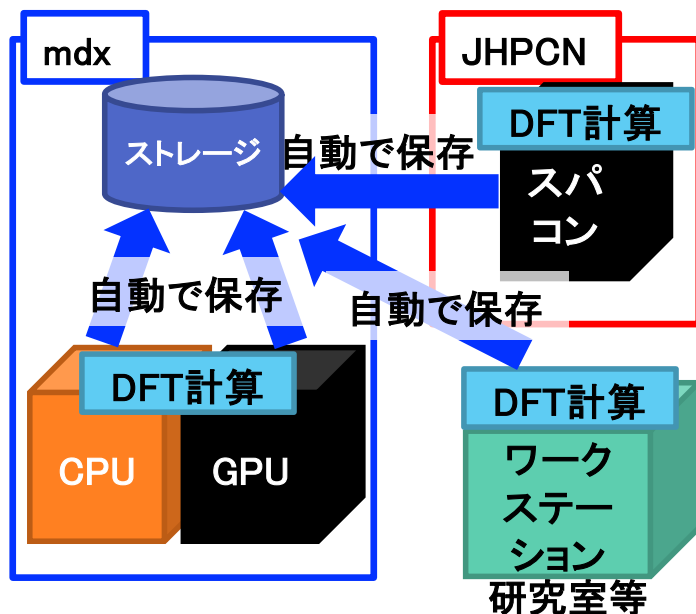


教師データを提供

- 機械学習ポテンシャルの自動生成
 - 将来的にはオンザフライでポテンシャルの自動改良
 - 足りない教師データを自動で取得(第一原理計算)

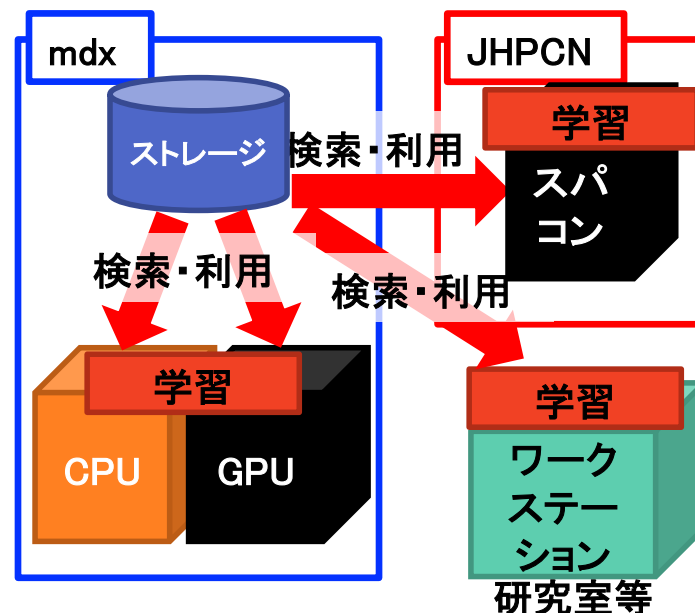
統合機械学習分子動力学システムの構築

教師データ収集イメージ



- 量子力学計算の結果を自動で保存
- データベースに自動で追加される

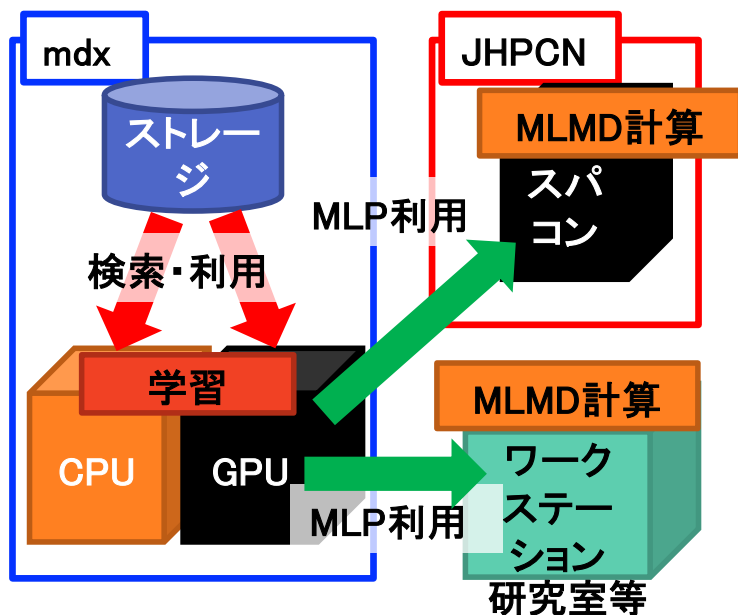
教師データ データベース 利用イメージ1



- 元素等で検索
- 様々な計算機環境で学習に利用

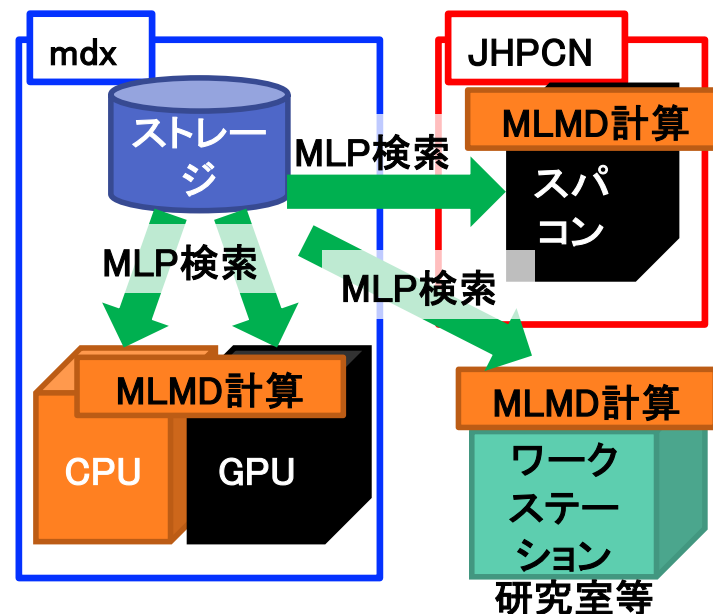
統合機械学習分子動力学システムの構築

教師データ データベース 利用イメージ2



- 大きな計算機資源が必要な学習はmdxで実施し、完成したMLPを様々な計算機環境でMLMDに利用

MLPデータベース 利用イメージ



- MLPのデータベースからMLPを取得し、mdxや様々な計算機環境でMLMDを実施

統合機械学習分子動力学システムの構築

・現状

- ・ 機械学習MDのための教師データの共有
- ・ 機械学習ポテンシャルの自動生成
 - ・ Pythonパッケージを作成中
 - ・ オープンソースの機械学習MDコード「n2p2」の自動実施コードを作成中

デモンストレーション

- ・ 今後の予定
 - ・ 他の機械学習MDコードも利用可能にする(aenet, ...)
 - ・ オンザフライ自己学習機の実装
 - ・ mds \Leftrightarrow スパコン: ジョブ投入、データ回収
 - ・ 教師データのデータベース作成
 - ・ Gakunin RDMとの連携

Gakunin RDMとの連携(予定)

- Gakunin RDM
 - <https://rdm.nii.ac.jp/>



Gakunin RDMとの連携(予定)

研究データ管理(RDM)サービス GakuNin RDMの概要



研究データ共有・管理機能



NII研究データ基盤や外部ツールとの連携



研究証跡の保存機能



時刻認証事業者のタイムスタンプで
ある時刻のファイルの存在を証明



機関利用のシステム管理者のための管理機能



組織内のRDMサービスのカスタマイズ

- 研究データ共有・管理機能は本プロジェクトの目的と同じ
- 一部Gakunin RDMの機能につなげたい
- 連携に向けてGakunin RDMの開発者に相談中
- まずはJupyter Notebookで連携

まとめ

- 機械学習分子動力学法におけるデータ共有
 - 共有すべきデータ: 第一原理計算の計算結果
 - 方法
 - 自動で機械学習ポテンシャルを作りつつ、データベースに加えていく
- 現状
 - 機械学習ポテンシャル生成コード(n2p2)の制御プログラムを作成
 - Gakunin RDMとの連携
- 今後
 - 機械学習ポテンシャル生成コードの制御プログラムの拡充、テスト、公開
 - 教師データのデータベース構築(個人 → グループ → 公開)
 - 機械学習ポテンシャルとは関係なく計算結果のデータベースとして有用かも

JHPCN課題 統合機械学習分子動力学システムの構築

共同研究者(敬称略)

- 東京大学: 鈴木豊太郎、渡邊聡、沖田泰良、華井雅俊、芝隼人、清水康司
- 名古屋大学: 横井達矢
- 九州大学: 加藤幸一郎
- 熊本大学: 島村孝平
- 名古屋工業大学: 小林亮
- 産業技術総合研究所: 安藤康伸
- 量子科学技術研究開発機構: 河野秀俊、石田恒
- 産業技術短期大学: 森英喜
- 大阪国際工科専門職大学: 富谷昭夫
- AGC株式会社: 浦田新吾、今村穰、吉田拓未
- 原子力機構: 永井佑紀、中村博樹、山口瑛子、小林恵太、板倉充洋、町田昌彦

ご興味をお持ち頂けましたら、お気軽にご連絡ください。

okumura.masahiko@jaea.go.jp

ご清聴ありがとうございました！