

京都大学	博士 (工学)	氏名	西山隆之
論文題目	Application of machine learning potential to predict grain boundary properties and development of its performant implementation (機械学習原子間ポテンシャルの結晶粒界構造探索への応用と高速化手法開発)		
<p>(論文内容の要旨)</p> <p>本論文は、分子動力学計算などに基づいて材料の性質を議論する上で重要となる、機械学習を利用した高精度原子間ポテンシャル（機械学習ポテンシャル）に関して、予測能力を検証し、高速な実行の方法を検討した結果をまとめたものであり、4章からなっている。</p> <p>第1章は序論であり、研究の目的と背景について一般的な特徴をまとめ、論文全体の流れを総括している。近年、電子構造に基づく第一原理計算が材料科学に広く利用されるようになり、様々な材料特性が定量的に予測可能になってきている。しかし、第一原理計算は高い計算コストを要求するため、幅広い空間・時間スケールに渡る材料科学の問題を全て第一原理計算することは現実的ではない。このような第一原理計算では困難な規模の材料科学の問題に取り組むために、経験的ポテンシャルと古典力学の組み合わせによる分子動力学計算などが行われてきた。経験的ポテンシャルは化学結合を物理モデル化し、物理特性を再現する目的で作成されたものである。この方法では第一原理計算と比較して大規模・長時間の計算が可能であるものの、想定した化学結合モデルから逸脱する構造に対して、正確な計算が困難になるという問題があった。そこで、化学結合に基づく物理特性の推定に代わって、原子周辺環境から構造特徴量を計算し、機械学習的手法を用いてポテンシャルエネルギー表面（PES）を推定する、機械学習ポテンシャルが開発された。一般に機械学習ポテンシャルは、きわめて幅広い構造に対して第一原理計算に準ずる精度でのPES推定が可能であるが、一方で訓練データや構造特徴量の選択が機械学習ポテンシャルの性能に決定的な影響を与えることが知られている。また、これまでの機械学習ポテンシャルの研究は手法開発に関するものが大多数を占めるが、近年ようやく挑戦的な応用が可能になってきた段階である。それに伴い、経験的ポテンシャルと比較して高い計算コストが、機械学習ポテンシャルを応用する際の問題の一つとして認識され始めている。本研究では、所属するグループにおいて第一原理計算の結果をもとに開発された機械学習ポテンシャルに関して、それを用いた大域的な結晶粒界構造探索と、グラフィックス・プロセッシング・ユニット（GPU）を用いた場合の高速化手法開発を議論している。結晶粒界構造探索の過程で、系統的に生成された一連のポテンシャルに対する粒界エネルギーの振る舞いを観察することで、機械学習ポテンシャルが適切に構成されていることを評価している。経験的ポテンシャルと機械学習ポテンシャルのそれぞれについて、結晶粒界構造探索の結果を比較している。高速化手法を適用したGPU実装に関しては、様々な粒子数およびプロセッサ並列数に対する実行速度を評価した結果を報告している。</p>			

京都大学	博士 (工学)	氏名	西山隆之
<p>第2章では、機械学習ポテンシャルを用いて、面心立方構造の単体金属 Ag, Al, Au, Cu, Pd, および Pt において、$\langle 100 \rangle$対称傾角粒界、$\langle 110 \rangle$対称傾角粒界、$\langle 100 \rangle$単純ねじれ粒界に関して、粒界構造および粒界エネルギーの角度依存性を評価した結果を報告している。最安定な結晶粒界構造は、それぞれの粒界構造モデルに関して、様々な剛体変位を加えることにより大域的に探索している。ここで使用した機械学習ポテンシャルの訓練データは、既知の基本的結晶構造を乱数で拡大縮小・変形させた構造から第一原理計算により生成されている。この訓練データは粒界構造データを一切用いずに生成されたものであるにもかかわらず、プレート最適上の全てのポテンシャルを用いて粒界エネルギーを計算した結果が第一原理計算の値に必要精度内に収束したことを報告している。さらに、経験的原子間ポテンシャルと機械学習ポテンシャルの両方で同じ大域的最適化手法を実行し、粒界エネルギーの予測値を比較したところ、粒界エネルギーが大きい構造に対して、経験的ポテンシャルでは第一原理計算の値から大きく外れる場合があった一方で、機械学習ポテンシャルでは第一原理計算の結果を良好に再現した。この計算結果は、本研究での機械学習ポテンシャルの訓練データ、構造特徴量および機械学習モデルの選択が適切であり、機械学習ポテンシャルは訓練データに含まれない欠陥構造に対しても予測能力が高いことを示している。</p> <p>第3章では、機械学習ポテンシャルの計算コストを改善する目的で、計算領域の空間分割に基づいて GPU を利用し、計算を高速化する手法を議論している。高速化手法開発は、ポテンシャル計算の繰り返し処理の高速化と空間分割の適用の二段階に分かれており、ポテンシャル計算の構造を整理しアルゴリズムの形でまとめている。アルゴリズムに現れる繰り返し処理のうち、部分系内部で閉じた演算は、プロセッサユニット内の並列処理により高速化している。空間分割は、機械学習ポテンシャルを計算するための特徴的な境界条件を用い、プロセッサ間の通信を大幅に削減することで実現している。その結果、空間分割に基づいて並列させたプロセッサユニットを用いて機械学習ポテンシャルを高速に計算することが可能になり、並列 GPU を用いた場合、CPU コアによる計算と比較して最大で約千倍の高速化を達成したことを報告している。また、単一の GPU ではメモリが不足し、機械学習ポテンシャルを用いた計算が不可能となるような大規模な問題に対しても、複数の GPU を利用することで機械学習ポテンシャルの高速計算が可能となることが確認された。</p> <p>第4章は結論であり、本論文で得られた成果について要約している。</p>			

(論文審査の結果の要旨)

本論文は、機械学習ポテンシャルの予測能力に関して結晶粒界構造を例にして検証し、また機械学習ポテンシャルを用いた大規模な分子動力学計算の高速な実行方法を提案したものであり、主な成果は以下のとおりである。

1. 面心立方構造の単体金属 Ag, Al, Au, Cu, Pd, および Pt において、 $\langle 100 \rangle$ 対称傾角粒界、 $\langle 110 \rangle$ 対称傾角粒界、 $\langle 100 \rangle$ 単純ねじれ粒界に関して、粒界構造および粒界エネルギーの角度依存性を評価した結果、機械学習ポテンシャルが第一原理計算の結果を再現したことから、機械学習ポテンシャルは訓練データに含まれない欠陥構造に対しても予測能力が高いことを示した。
2. シミュレーション領域の空間分割に基づく並列化を用いて機械学習ポテンシャルを高速化する方法を新たに開発した。並列 GPU を用いた場合、単一の CPU コアによる実行速度と比較して約千倍の高速化を達成した。また単一の GPU ではメモリが不足して機械学習ポテンシャルによる計算を実行できない規模の系に対しても、複数の GPU を並列してシミュレーション領域を分割することで計算が可能になることを示した。

これらの結果から、機械学習ポテンシャルは訓練データに含まれない未知構造の探索に使用可能であることが示され、また新しく開発された機械学習ポテンシャルの GPU 並列実装により、高精度な機械学習ポテンシャルを用いた大規模計算が可能になることが示された。その成果は、学術上、實際上寄与するところが少なくない。

よって本論文は博士（工学）の学位論文として価値あるものと認める。また、令和4年2月15日、論文内容とそれに関連した事項について試問を行って、申請者が博士後期課程学位取得基準を満たしていることを確認し、合格と認めた。