

## 科学研究費助成事業 研究成果報告書

平成 29 年 6 月 28 日現在

機関番号：33503

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26800203

研究課題名(和文) 帰納法アプローチによるシミュレーションデータからの物性予測と解釈

研究課題名(英文) Prediction and interpretation of physical properties from simulation data by induction approach

研究代表者

杉山 歩 (Sugiyama, Ayumu)

山梨英和大学・人間文化学部・准教授

研究者番号：20586606

交付決定額(研究期間全体)：(直接経費) 1,500,000円

研究成果の概要(和文)：本研究はマテリアルズインフォマティクスにより、金属合金の物性予測と結果の解釈から新たな物理的知見を得るための方法論の確立を目的として行った。本研究では金属錯体及び2元金属合金の第一原理計算と既知の実験データから構築した物性データベースを構築し、それぞれのデータ群に対する機械学習・データマイニングにより教師なし学習による分類を行った。分類されたデータ群から物理的意味の解釈を行い、本研究で対象とした金属錯体に関するデータ群、2元金属合金に関するデータ群共に分類されたデータ群には既知の物性による分類と相違無く、機械学習による分類が従来の物性知識を意味論と立場からも正しく分類出来る事を示した。

研究成果の概要(英文)：The purpose of this research is establishing a methodology for obtaining new physical knowledge from prediction of physical properties of metal alloys and interpretation of results by Materials Informatics using machine learning and data mining techniques. At first, physical property database is constructed from the first principle calculation of metal complex and binary metal alloy and known experimental data and classified by unsupervised learning by machine learning and data mining for each data group. We also interpreted the physical meaning from the classified data group and examined the semantics of each classified data group. In the data group classified as data group on the metal complex and the data group on the binary metal alloy covered in this research, classification by machine learning is the same as classification by known physical properties, and semantics of the conventional physical property knowledge are also indicated.

研究分野：計算物理学

キーワード：データマイニング 第一原理計算 マテリアルズインフォマティクス

## 1. 研究開始当初の背景

研究開発当初、マテリアルサイエンス分野の研究開発は計算機の高速化と低コスト化から計算機シミュレーションは広く一般に普及しており、「実験結果の電子論レベルにおける理解・解析」や「実験を行う事無く電子軌道や構造推定」など実験や理論との協力関係のもとに多大な成果を上げていた。これらの成果は学術研究分野のみならず材料開発を行う産業界からも強い関心を集めており、材料科学分野では物性予測にとって不可欠なツールと語られるまでになっていた。この材料分野研究における計算機の新しい活用法として「目標機能の設計手法の確立」を目的としたインフォマティクスの利活用に関する検討が米国を中心に先進各国において始まっており、我が国が今後も科学技術立国としての地位を維持し、産業競争力における優位を維持するためにも、本分野の確立と優位性の確保は不可欠であるとの認識が広がり始めている段階であった。

また、米国では情報学による材料科学の推進を「マテリアルズ・インフォマティクス」と称して今後の重点成長分野と位置づけ、2012年に国家プロジェクト「Materials Genome Initiative」計画をスタートさせており、我が国においてもマテリアルズ・インフォマティクスを「計算機科学(データ科学、計算科学)と物質・材料の物理的・科学的性質に関する多様で膨大なデータとを駆使して、物質・材料科学の諸問題を解明するための科学的技術的手法」と定義し、その重要性についての議論が始まっていた。特にマテリアルズ・インフォマティクスの手法論によりこれまで10~30年の期間を要していた新規先端素材の開発期間を数年にまで縮める事が可能であるとの見方を持ち始めていた。

## 2. 研究の目的

本研究では「第一原理シミュレーションデータを利用したマテリアルズ・インフォマテ

イクスの枠組みの構築と合金の融点予測」を課題に設定し、シミュレーションとそのデータの新しい活用手法の提案を行う。具体的研究目標は以下の通り。

### a. シミュレーションデータからの各種物性予測

二元合金に関する情報は結晶構造を始めとし融点、沸点、電子移動度など物性ハンドブック等によって多くの物性情報を知る事が出来る。本研究ではこれら既知の基本物性情報の予測をシミュレーションデータに対する回帰分析を行い、シミュレーションデータに基づいたマテリアルズ・インフォマティクスの実証性についての検証を行う。

### b. 物質間の関連性・相関性の定量的理解

予測に使用した線形回帰モデルを用いて有向グラフを作成する。特定機能を有する材料を設計する上で必要となる情報を合理的に探索し(逆問題的解法)物質間の関連性を説明する。

## 3. 研究の方法

マテリアルズ・インフォマティクスには機械学習・データマイニングの方法論に則り図4に示すフローチャートに基づいた研究を進める予定である。以下にそれぞれのステージでの具体的な工夫について説明する。

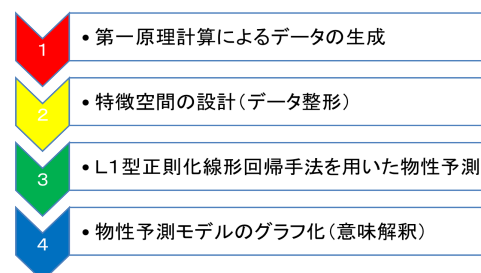


図4：研究のフローチャート

### (1) データ生成:

データの生成には密度汎関数理論に基づいた第一原理計算により行う。計算に使用するソフトウェアはDmol3およびGaussian09プログラムパッケージを予定している。両ソフトウェアは北陸先端科学技術大学院大学所

有の大型計算機に導入済みであり、申請者は既に十分な利用経験を有している。本シミュレーションでは全ての金属種を分子の形で可能な限りシンプルな構造として表現する事で計算量の削減に努める。

#### (2)特徴空間の設計(データ整形):

マテリアルズ・インフォマティクスを始め、機械学習・データマイニングにおいて最も重要なプロセスが前処理としての特徴空間の設計である。本課題の様な物理データにおいては物理・材料科学分野での専門的な知識に基づいた特徴空間の設計が必要となる。申請者は物理化学分野でのシミュレーション経験を十分に積んでおり、これまでの経験と専門家のアドバイスのもと、特徴空間を設計し、物質データベースを作成する。必要とされる特徴空間として最もシンプルなものでは結合距離や電荷など既知の基本情報があげられる。このほかにも必要に応じて様々な物理パラメータを計算し、次元数を増やしてゆく。データ処理の観点からも極端に大きな次元数は必要とされないため、30~50次元程度の特徴空間を設計する予定である。

#### (3)L1 型正則化線形回帰手法を用いた物性予測:

上述(1),(2)により作成したシミュレーションデータベースを元に以下の物性予測を実行する。

具体的には以下の2つのモデルを想定し、それぞれの物性予測を実行する。

二元合金の各種物性予測

#### 4. 研究成果

##### A. 金属錯体の分類

本研究ではマテリアルズインフォマティクスにより膨大な金属の組み合わせの中から関連性を発見する手法開発を目的としている。研究に使用したデータベースの構築にはABO型パイロクロア構造のうちA, B原子を核としたパイロクロア型金属錯体(6配位)分子48種(A-site:29種, B-site:19種)に対して

第一原理計算を行い、その計算結果から中心金属原子と配位原子のうち平面方向の1原子の電荷、結合長、その他物理パラメータによるデータベースの構築を行った。また、既知のデータベースからも同様に配位結合長のデータの追加を行った。構築されたデータベースから教師無し学習法による分類により、ABO 錯体を大きく2つのタイプへの分類に成功した。この2つの分類の物理的特徴に関する考察を行い、この分類がシャロンイオン半径と密接な関連性がある事が確認された。

##### B. 2元合金のクラスター分析

本研究では2つの元素からなる合金(2元金属合金)既知のデータ群(結合長、融点、沸点など)と第一原理計算から得られたデータ群を用いた物性予測を行った。本研究では230種の2元金属合金に対してK-meansクラスタリングによりグループ分けを行った。対象としたデータはvolume, Formation\_energy, Band\_gap, A\_Z, A\_e\_negativity, A\_valence\_e, A\_first\_ionization, A\_boiling\_point, A\_melting\_point, A\_atomic\_radius, A\_average\_ionic\_radius, B\_Z, B\_e\_negativity, B\_valence\_e, B\_first\_ionization, B\_boiling\_point, B\_melting\_point, B\_atomic\_radius, B\_average\_ionic\_radiusの19パラメータであり、クラスタリングにより数種類の分類を試行した結果、クラスター数3の分類結果がもっともクラスター毎の特徴の解釈が行いやすい結果である事が示された。これら3つに分類されたクラスターは1.ホウ素・炭素合金類、2.遷移金属合金類、3.典型金属合金類となっており4クラスターでの分類では典型金属元素がより細かく分類され、5クラスターでの分類ではさらに遷移金属がより細かく分類されるという結果となった。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 0件)

〔学会発表〕(計 1件)

“Structural stability of bismuthniomium oxide films in solution process”  
Ayumu Sugiyama, Tadaoki Mitani, Akira Goto  
The international chemical congress of pacific basin societies 2015 (Hawaii, USA)

〔図書〕(計 0件)

〔産業財産権〕

出願状況(計 0件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年月日：  
国内外の別：

取得状況(計 0件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
取得年月日：  
国内外の別：

〔その他〕  
ホームページ等

## 6. 研究組織

### (1) 研究代表者

杉山歩 (Sugiyama Ayumu)  
山梨英和大学・人間文化学部・准教授  
研究者番号：20586606

### (2) 研究分担者

なし

### (3) 連携研究者

なし

### (4) 研究協力者

ダム ヒョウ チ (Dam Hieu Chi)  
北陸先端科学技術大学院大学・知識マネジメント領域・准教授  
研究者番号：70397230