

第一原理計算ソフトウェア Advance/PHASE の MI 機能

胡 春平* 岡崎 一行**

Tools for Materials Informatics in Advance/PHASE

Chunping Hu* and Kazuyuki Okazaki**

第一原理計算ソフトウェア Advance/PHASE Ver. 4.0 に導入したマテリアルズ・インフォマティクス(MI)機能について紹介する。

Keywords: マテリアルズ・インフォマティクス、材料データベース、データマイニング、機械学習

1. はじめに

データマイニングや機械学習のような情報科学技術を材料分野に応用したマテリアルズ・インフォマティクス(MI)と呼ばれる概念が注目されてきている。従来は研究者の勘や経験に基づいて材料研究・開発は行われてきたが、開発期間の短縮や開発コストの削減などを目指して MI を活用しようという機運が高まってきている。これだけでなく、MI はデータベースから目的とする材料を直接検索することによって、材料ユーザーのニーズに合わせた開発が効率的に行える可能性や従来の手法ではたどり着くことができなかった未知の材料の発見につながる可能性があると考えられている。

機械学習は統計数理に基づいた技術で、人工知能(AI)の中核をなすものである。AI といわれると、人間の能力を凌駕し、AI や機械学習があればなんでも解決できると思われがちだが、研究者が物理・化学・材料学などの専門知識を用いて目標やルールを設定し、AI や機械学習をサポートツールとして適切に用いて材料開発を行っていくことが MI の本質であると考えられる。

MI には、材料のさまざまな物性値のビックデータが必要となる。実験でそのようなデータを

作成する手法のひとつとして、コンビナトリアル実験[1, 2]がある。また、計算機シミュレーションを用いてデータを作成することができる。第一原理計算は、物理学の第一原理（古典力学、電磁気学、統計力学、量子力学）だけにに基づき、原子・電子レベルで材料特性をシミュレーションする手法であり、MI のビックデータを作成するツールとして用いることができる。そこで、第一原理計算ソフトウェア Advance/PHASE Ver. 4.0 では PHASE-Viewer に MI に関連するいくつかの機能を追加したので、その機能を紹介する。また、第一原理計算について詳細に知りたい方は、参考文献[3]を参照されたい。

2. Advance/PHASE の MI 機能

Advance/PHASE の MI 機能は、オープンソースである pymatgen[4]や matminer[5]との連携で実現している。

2.1. データベース検索

Advance/PHASE Ver. 4.0 では、公開データベースである Materials Project[6]と Citrine[7]を PHASE-Viewer から利用できる。このとき、

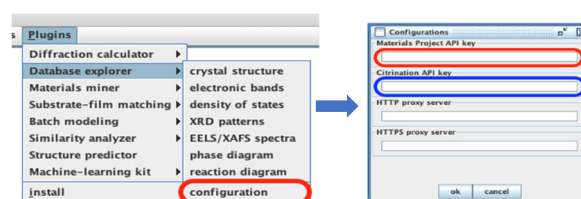


図 1 PHASE-Viewer での API Key の設定画面

*アドバンスソフト株式会社 第1事業部
1st Computational Science and Engineering
Group, AdvanceSoft Corporation

**アドバンスソフト株式会社 関西支店
Kansai Branch, AdvanceSoft Corporation
アドバンスシミュレーション 2020.9 Vol.28

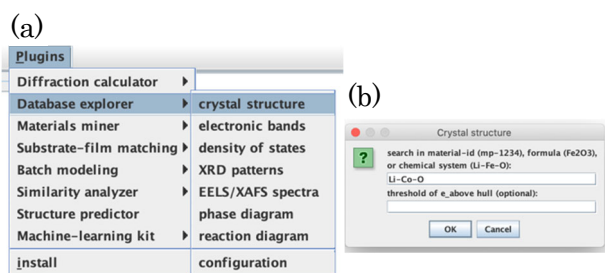


図 2 データベース検索をするときの操作(a)と検索する材料情報の設定パネル(b)。

PHASE-Viewer で API Key の設定(図 1)が必要である。Materials Project では、ログイン(無料のユーザー登録あるいは Google アカウントが必要です)後に Dashboard をクリックすると API Key が表示される。Citrine では、ログイン(無料のユーザー登録あるいは Google アカウントが必要です)後にユーザー名>Account settings を選択すると API Key が表示される。データベースを利用せず、ローカルの情報のみを利用する場合は API Key の設定などは必要ない。

Advance/PHASE Ver. 4.0 で検索可能なデータは、

- (a) 結晶構造[crystal structure]
- (b) バンド構造[electronic bands]
- (c) 状態密度[density of states]
- (d) XRD スペクトル[XRD patterns]
- (e) EELS/XAFS スペクトル[EELS/XAFS spectra]
- (f) 相図[phase diagram]
- (g) 反応ダイアグラム[reaction diagram]

となっている。

結晶構造の検索では、Plugins>Database explorer>crystal structure の順に選択(図 2(a))する。図 2(b)の検索する材料情報の設定パネルに、Materials Project での ID や組成式、化学種を入力し、OK ボタンを押すことにより、条件に一致する構造ファイルを各フレームに取り込んだ原子配置ビューアーが起動する。このとき、構造安

```
INFO -
INFO - ----- start of crystal structure search for Li-Co-O -----
INFO -
INFO - structure found: Li24 Co4 O16 materials id: mp-18925 e_above_hull: 0 index: 0
INFO - structure found: Li9 Co7 O16 materials id: mp-1175419 e_above_hull: 0 index: 1
INFO - structure found: Li20 Co8 O18 materials id: mp-1185179 e_above_hull: 0 index: 2
INFO - structure found: Li8 Co16 O32 materials id: mp-774082 e_above_hull: 0 index: 3
INFO - structure found: Li47 Co8 O32 materials id: mp-1177487 e_above_hull: 0 index: 4
INFO - structure found: Li4 Co4 O8 materials id: mp-949273 e_above_hull: 0 index: 5
INFO - structure found: Li8 Co4 O12 materials id: mp-1173879 e_above_hull: 0 index: 6
INFO -
INFO - - totally 7 structure(s) found with lowest e_above_hull = 0
INFO -
INFO - ----- end of crystal structure search for Li-Co-O -----
```

図 3 Li-Co-O のデータベース検索結果のログ

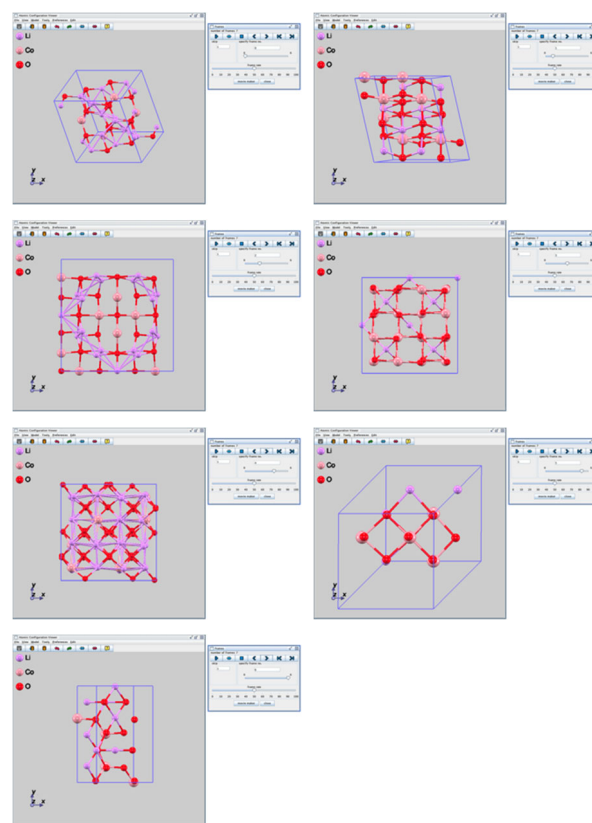


図 4 Li-Co-O のデータベース検索で得られた結晶構造

定性も任意で入力できる。図 2(b)に示したように、Li-Co-O と入力してみると、Materials Project の検索が始まり、7 個の結晶構造データが見つかった(図 3)ことが分かる。この場合、原子配置ビューアーのフレーム 0 から 6 に構造を表示する(図 4)。原子配置ビューアーで保存操作を行うと、現在のフレーム、全てのフレーム、特定のフレームのいずれかを選択(図 5)して、Advance/PHASE を用いた第一原理計算が実行できるようになる。このとき、all frames や specified frames を選択すると、複数のプロジェクトを一括で作成できる。ただし、計算に必要な初期設定は全てデフォルト値になっているので、精度などを考えるときは、ユーザーによる個別の変更(あるいは一括変更)が必要となる。

(b)から(g)の検索機能についても、図 2(a)でどのデータを検索するかを選択し、結晶構造の場合と同様に、利用したい材料の情報やその他の必要な情報を適宜入力することによって、データベース上の情報を取得することができる。

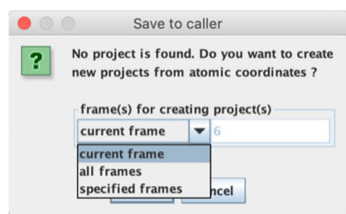


図 5 Advacen/PHASE 計算用のプロジェクトを作成するための構造選択のダイアログ

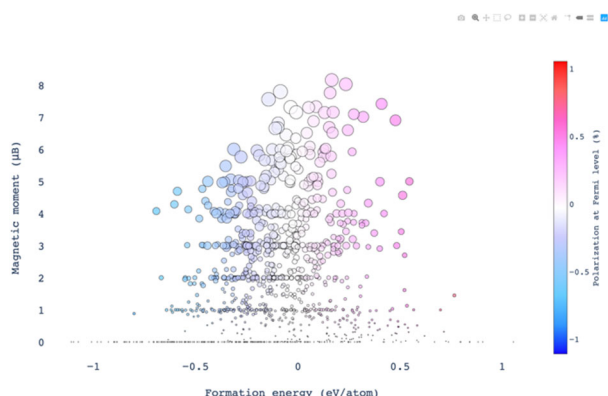


図 6 磁性材料のデータをプロットした結果

2.2. データマイニング

Advance/PHASE Ver. 4.0 で検索可能なデータは、

- (a) 2 次元材料[2D materials]
- (b) 鉄鋼材料[steel materials]
- (c) 弾性材料[elastic materials]
- (d) 磁性材料[magnetic materials]
- (e) 誘電材料[dielectric materials]
- (f) 圧電材料[piezoelectric materials]
- (g) 熱電材料[thermoelectric materials]
- (h) ユーザー定義[select materials dataset]

となっている。

磁性材料のマイニングでは、Plugins>Materials miner>magnetic materials の順に選択すると、自動的にデータベースから情報を取得し、図 6 のように横軸に形成エネルギー、縦軸に磁気モーメント、色でフェルミ準位での分極率を表している。このプロットでは、プロットされた丸印の上にマウスオーバーすると、材料の情報が表示される。(a)から(g)の他の機能についても、それぞれの材料に関して考えるべき物性値がプロット

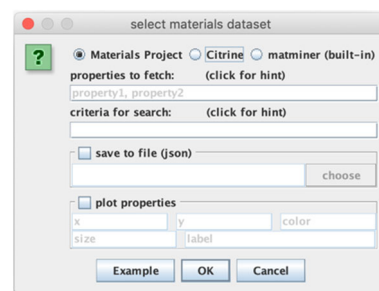


図 7 データベースの選択と材料の検索条件を設定するダイアログ

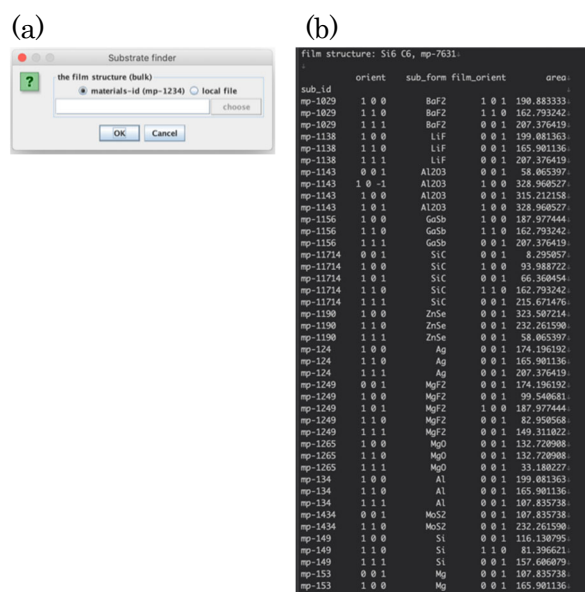


図 8 基板探索のダイアログ(a)と検索結果(b)

される。

(h) ユーザー定義では、Plugins>Materials miner>select materials dataset の順に選択すると、図 7 のようなダイアログが現れ、利用するデータセットを選択し、収集する物性値やプロット条件を入力することでデータマイニングを行うことができる。

2.3. 基盤/膜マッチング

Advance/PHASE Ver. 4.0 で利用可能な機能は、(a) 基板探索[substrate finder]

(b) 界面ビルダー[interface builder]

となっている。

基板探索では、Plugins>Substrate-film matching>substrate finder の順に選択し、図 8(a) のようなダイアログで成膜に用いたい材料の Materials Project の ID を入力する。結果は、例

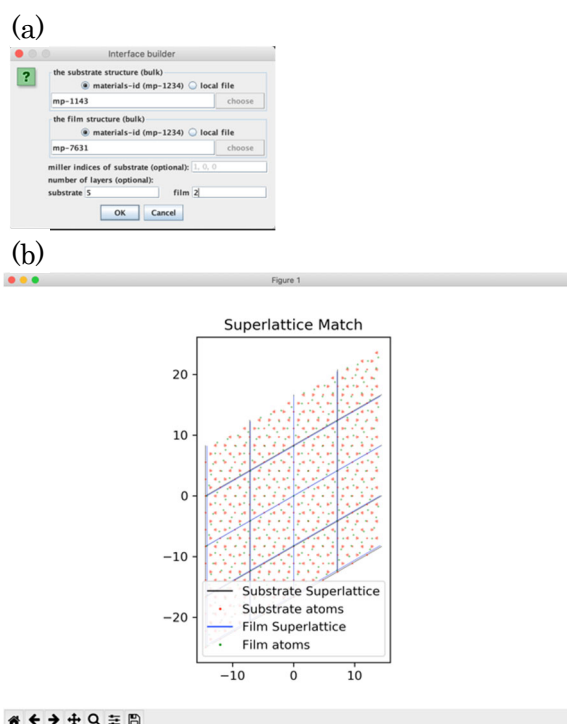


図 9 界面ビルダーの設定ダイアログ(a)と超格子の整合性の表示(b)

として 6H-SiC の Materials Project の ID(mp-7631)を入力すると、図 8(b)のように整合する材料とその時の基板と成膜材料の面指数が一覧として表示 (図 8(b)は検索結果の一部である) される。

界面ビルダーでは、Plugins>Substrate-film matching>interface builder の順に選択し、図 9(a)のようなダイアログで基板材料と膜材料を入力する(例では、mp-7631 と基板探索で検出された材料の mp-1143 とした)。次に、図 9(b)に示した超格子としての整合性が表示される(スラブの候補ごとに複数表示される場合がある。今回の例では、4 個の候補が表示される)。この画面を閉じると、原子配置ビューアーが起動し、候補構造が各フレームに取り込まれている。今回の例では、36 個の界面モデルが作成できた。結晶構造データベースの検索の時と同様に、保存操作を行うと、現在のフレーム、全てのフレーム、特定のフレームのいずれかを選択 (図 5) して、Advance/PHASE を用いた第一原理計算が実行できるようになる。このとき、all frames や specified frames を選択すると、複数のプロジェクトを一括で作成できる。ただし、計算に必要な初期設定は全てデフォルト値になっているので、精度などを考える

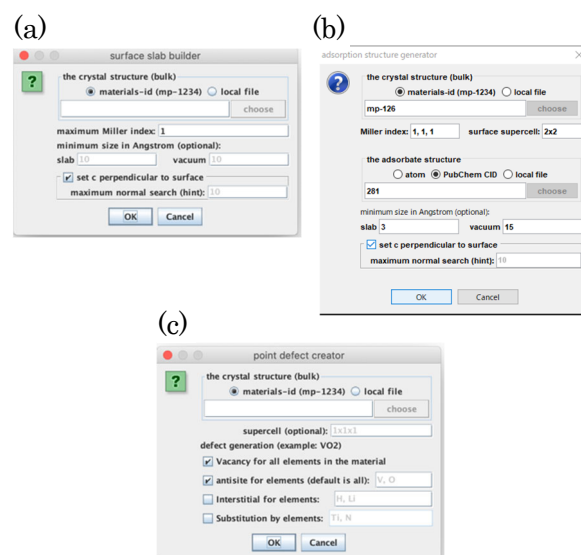


図 10 バッチモデリングの設定ダイアログ。(a) 表面構造作成、(b) 表面吸着構造作成、(c) 点欠陥作成。

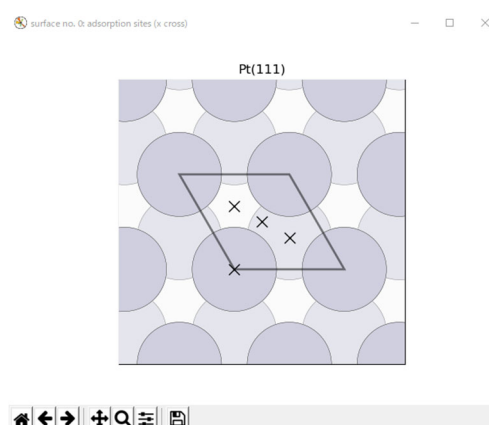


図 11 スラブの候補と吸着位置の候補

ときは、ユーザーによる個別の変更 (あるいは一括変更) が必要となる。

2.4. バッチモデリング

Advance/PHASE Ver. 4.0 で利用可能な機能は、
(a) 表面構造作成 [surface slab builder]
(b) 表面吸着構造作成 [adsorption structure generator]
(c) 点欠陥作成 [point defect creator]
となっている。

表面構造作成では、Plugins>Batch modeling>surface slab builder の順に選択し、図 10(a)のようなダイアログで、スラブに変換したい材料、スラブ化において考慮するミラー指数の最大値を

入力する。作成できた構造ファイルを各フレームに取り込んだ原子配置ビューアーが起動する。フレームの切り替えやプロジェクトの作成は、これまでと同様に、保存操作を行うと、現在のフレーム、全てのフレーム、特定のフレームのいずれかを選択(図 5)して、Advance/PHASE を用いた第一原理計算が実行できるようになる。このとき、**all frames** や **specified frames** を選択すると、複数のプロジェクトを一括で作成できる。ただし、計算に必要な初期設定は全てデフォルト値になっているので、精度などを考えるときは、ユーザーによる個別の変更（あるいは一括変更）が必要となる。

表面吸着構造作成では、**Plugins>Batch modeling>adsorption structure generator** の順に選択し、図 10(b)のようなダイアログでスラブに変換したい材料(図では Pt(mp-126))、表面のミラー指数(1 1 1表面)、表面セルの超格子の大きさ(2×2)、吸着分子の構造(CO(PubChemID 281))などを入力する。その後、図 11 のように作成されたスラブの候補と吸着位置の候補が表示される(スラブの候補ごとに複数表示される場合がある)。この画面を消すと、作成できた構造ファイルを各フレームに取り込んだ原子配置ビューアーが起動する。フレームの切り替えやプロジェクトの作成は、これまでと同様に、保存操作を行うと、現在のフレーム、全てのフレーム、特定のフレームのいずれかを選択(図 5)して、Advance/PHASE を用いた第一原理計算が実行できるようになる。このとき、**all frames** や **specified frames** を選択すると、複数のプロジェクトを一括で作成できる。ただし、計算に必要な初期設定は全てデフォルト値になっているので、精度などを考えるときは、ユーザーによる個別の変更（あるいは一括変更）が必要となる。

点欠陥構造作成では、**Plugins>Batch modeling>point defect creator** の順に選択し、図 10(c)のようなダイアログで母材の材料、作成する点欠陥の条件(空孔生成、アンチサイト生成、不純物置換、不純物挿入)を入力する。作成できた構造ファイルを各フレームに取り込んだ原子配置ビュー

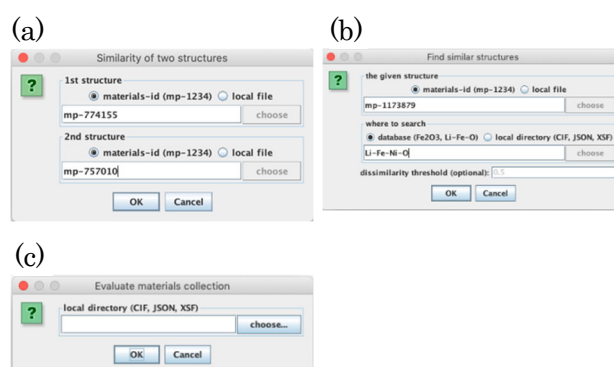


図 12 類似性分析の設定ダイアログ。(a) 2 つの構造比較、(b)似た構造の探索、(c)一連の構造における類似性の評価。

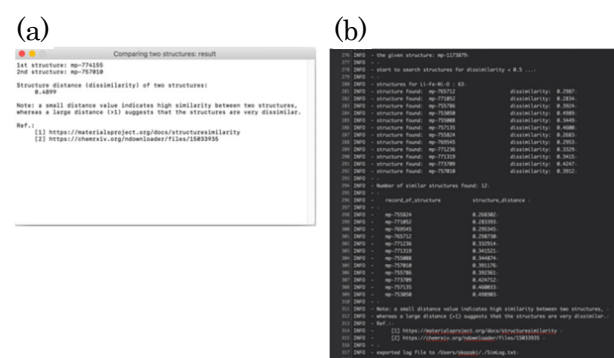


図 13 類似性分析の結果。(a)2 つの構造比較、(b)似た構造の探索。

アーが起動する。フレームの切り替えやプロジェクトの作成は、これまでと同様に、保存操作を行うと、現在のフレーム、全てのフレーム、特定のフレームのいずれかを選択(図 5)して、Advance/PHASE を用いた第一原理計算が実行できるようになる。このとき、**all frames** や **specified frames** を選択すると、複数のプロジェクトを一括で作成できる。ただし、計算に必要な初期設定は全てデフォルト値になっているので、精度などを考えるときは、ユーザーによる個別の変更（あるいは一括変更）が必要となる。

バッチモデリングでは、考慮すべきモデルが自動で作成されるので、経験が少なくても、計算用のモデルを作成することが可能である。しかし、自動作成されたモデルが、物理的に意味あるモデルであるかはユーザーが判断する必要がある。

2.5. 類似性分析

Advance/PHASE Ver. 4.0 で利用可能な機能は、

(a) 2 つの構造の比較[compare two structure]

(b) 似た構造の探索[find similar structures]

(c) 一連の構造における類似性の評価[evaluate structure collection]
となっている。

2 つの構造の比較では、Plugins>Similarity analyzer>compare two structures の順に選択する。図 12(a)のようなダイアログで比較する 2 つの構造の情報(例では、 $\text{Li}_2\text{FeO}_3(\text{mp-774155})$ と $\text{Li}_3\text{Fe}(\text{NiO}_2)_4(\text{mp-757010})$)を入力する。結果は、図 13(a)のように類似度が表示される。類似度が 1 より大きな材料はあまり似ていないことを示している。

似た構造の探索では、Plugins>Similarity analyzer>find similar structures の順に選択する。図 12(b)のようなダイアログで比較の元となる材料(例では $\text{LiCoO}_2(\text{mp-1173189})$)の情報とどのデータを探索するか(例では Li-Fe-Ni-O を含む系)を入力する。結果は、図 13(b)のようなログが表示され、別ウィンドウでは設定した類似度(デフォルトでは 0.5)以下となる情報だけが表示される。

一連の構造における類似性の評価では、Plugins>Similarity analyzer>evaluate structure collection の順に選択する。図 12(c)のようなダイアログで、比較する一連の構造ファイルが格納されているフォルダを入力する。結果は、構造全体での類似度に関する統計データ(最小値、最大値、平均値、偏差値)が表示される。

2.6. 構造予測

構造予測の機能は、Plugins>Structure predictor の順に選択する。図 14 のようなダイアログで、予測に使うイオン種と電荷の情報(図 14 の例では (Ba, Ca)-Ti-O)を入力する。存在確率の閾値も制御できる。予測により得られた構造を描くフレームに取り込んだ原子配置ビューアーが起動

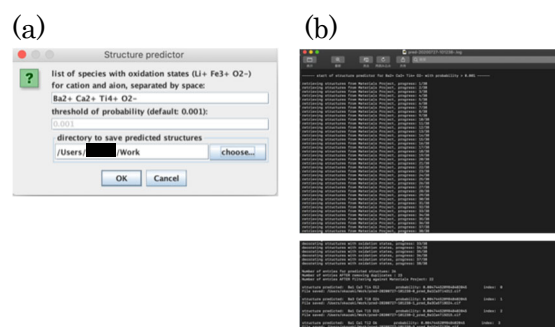


図 14 構造予測の設定ダイアログ(a)と結果のログファイルの一部(b)。

する。フレームの切り替えやプロジェクトの作成は、これまでと同様に、保存操作を行うと、現在のフレーム、全てのフレーム、特定のフレームのいずれかを選択(図 5)して、Advance/PHASE を用いた第一原理計算が実行できるようになる。このとき、all frames や specified frames を選択すると、複数のプロジェクトを一括で作成できる。ただし、計算に必要な初期設定は全てデフォルト値になっているので、精度などを考えるときは、ユーザーによる個別の変更(あるいは一括変更)が必要となる。

2.7. 機械学習ツール

Advance/PHASE Ver. 4.0 で利用可能な機能は、
(a) 既存の学習済みモデルを用いた材料物性値の予測[using pre-built models]
(b) 構造と物性値を用いた機械学習モデルのトレーニング[training with structures]
(c) 学習済みのモデルを用いた材料物性値の予測[predicting target property]
となっている。利用できる機能や予測できる物性値は限定的だが、機械学習を用いた材料開発のツールを提供している。

3. まとめ

Advance/PHASE Ver. 4.0 の PHASE-Viewer に搭載した MI 機能について紹介した。材料開発において AI を利用するデータマイニングのような情報科学技術を利用する MI は、これからも発展していく分野であると考えられる。MI については、さまざまな教科書[8-10]が出版されているの

で、詳細はそちらを参考されたい。

第一原理計算ソフトウェア Advance/PHASE は、これからも公開されている MI ライブラリの利用や独自で MI 機能の開発を進めて、より使いやすいツールにしていくことを計画している。

※ 技術情報誌アドバンスシミュレーションは、アドバンスソフト株式会社 ホームページのシミュレーション図書館から、PDF ファイルがダウンロードできます。(ダウンロードしていただくには、アドバンス/シミュレーションフォーラム会員登録が必要です。)

参考文献

- [1] H. Koinuma, I. Takeuchi, Nat. Mater. 3 (2004) 429.
- [2] I. Takeuchi, O. O. Famodu, J. C. Read, M. A. Aronova, K.-S. Chang, C. Craciunescu, S. E. Lofland, M. Wuttig, F. C. Wellstood, L. Knauss, A. Orozco, Nat. Mater. 2 (2003) 180.
- [3] R. M. Martin, “Electronic Structure: Basic Theory and Practical Methods”, Cambridge University Press (2008). [訳：寺倉清之、寺倉郁子、善甫康成、「物質の電子状態」丸善出版 (2012)]
- [4] <https://pymatgen.org>
- [5] <https://hackingmaterials.lbl.gov/matminer/>
- [6] <https://materialsproject.org>
- [7] <https://citration.com/>
- [8] 岩崎悠真、「マテリアルズ・インフォマティクス 材料開発のための機械学習超入門」日刊工業出版社 (2019).
- [9] O. Isayev, A. Tropsha, S. Curtarolo, “Materials Informatics: Method, Tools, and Applications”, Wiley-VCH (2019).
- [10] 船津公人、柴山翔二郎、「実践マテリアルズインフォマティクス Python による材料設計のための機械学習」近代科学社 (2020).

Advance/PHASE は、文部科学省の国家プロジェクト「戦略的基盤ソフトウェアの開発」(<http://www.ciss.iis.u-tokyo.ac.jp/fsis/index.html>)とそれに続く「革新的シミュレーションソフトウェアの研究開発」(<http://www.ciss.iis.u-tokyo.ac.jp/rss21/index.html>)によって開発された第一原理計算ソフトウェアを、アドバンスソフト株式会社が独自に開発を進めてきた計算ソフトウェアです。