HPCS2012
2012/1/25

材料設計シミュレーションにおける自動規則抽出の試み

林　亮子[†]
†：金沢工業大学　情報学部　情報工学科　〒924-0838 石川県白山市八束穂３－１

近年では計算機が高速化し，グリッド／クラウドコンピューティング環境が整備されてきたため大量の計算資源が利用できる．さらに，多くの材料設計シミュレーションプログラムが開発されており，容易に大量の材料設計シミュレーションが行えるようになってきた．しかし，大量の材料設計シミュレーション結果は大量の数値データであり，シミュレーション結果の有効活用にはデータの自動処理が必要である．一方で大量のデータから知識発見を行うデータマイニング技術も近年著しく発展し，誰でも容易にデータマイニングが行えるようになってきている[1]．本ポスターでは，材料設計シミュレーション結果データを自動処理して知識を発見することを目標として，試験的に規則発見を行った結果を報告する．

本研究では広く使われる統計解析向けプログラミング言語であるR言語[2]を使用する．Rには近年多くのデータマイニングパッケージが実装され，無償公開されている．材料設計については，性質が良く知られていて計算が短時間で終了する水分子の構造最適化シミュレーションを最初の例題とする．原子の初期配置は乱数を用いて決定するが，初期配置と最適構造の間の距離と計算の繰り返し回数，さらに結果として最適構造が得られるかどうかとの間の規則を調べる．

定性的には，最適構造付近の初期配置であれば計算の繰り返し回数は小さくて最適構造が得られる確率が大きく，最適構造から離れるほど計算の繰り返し回数が増加して最適構造が得られる確率が小さい．そこで，最適構造が得られる初期配置の範囲を定量評価することを目標とする．データマイニング手法にもいろいろあり，この場合も複数の手法が適用可能であるが，最適構造が得られる範囲の規則は言語化できるため，連関規則，すなわち「Aが成立すればBが起こる」という形式での規則発見を試みる．これは条件付き確率 $P(B|A)$ が大きくなるようなAとBを探すことと同じである．当日は，R言語の連関規則パッケージarulesを利用した結果を報告する予定である．

References

[1] 「データマイニング入門」，豊田秀樹，東京図書，(2008).
[2] “R: A language and environment for statistical computing“, R Development Core Team, R Foundation for Statistical Computing, Vienna, (2009).

謝辞
本研究の一部は科学研究費補助金　基盤（C）課題番号23500138による．関係各位に感謝する．

ⓒ 2012 Information Processing Society of Japan