材料科学における知識管理システムの研究

徐一斌　衣川純一　八木晃一
（独）物質・材料研究機構　材料基盤情報ステーション

1.緒言

これまでの材料研究と産業活動で蓄積されてきた大量の材料データおよび経験・理論は、材料科学において貴重な知識であり、重要な開発基盤である。そのためそれらを取り扱う材料知識の管理・提供システムは、材料の選択・使用や、新材料の設計・開発をサポートする等、社会生活と科学研究についての重要な情報源となる。本研究は、材料データと材料知識に関するデータの特徴、応用ニーズ等を分析した上で、材料科学における知識管理システムのモデル構築、およびそれらを実現する技術を検討・提案を行うものである。

2. 材料科学における知識

材料科学における知識は、材料に関するデータと材料科学理論から構成されている。材料データとは、材料の構造や特性など、特定の材料に関する知識であり、それには実験で得たファクトデータと、数値シミュレーション等の手法で得た計算データの二種類がある。材料科学の理論とは、特定の材料に限られず、全て或いは一部の材料に適用できる知識であり、材料科学で使われる概念の定義と、材料の構成・特性に関しての原理、定理、ルール等が挙げられる。

これら材料科学の知識は、以下の特徴を有する。

（1）高度な専門性

材料科学は、金属やセラミックス、高分子、複合材料等といった様々な材料の種類により分類されるだけでなく、電子ディバイス材料や、建築材料、空間材料等の応用領域や、また材料の製作・測定方法といった研究手法等により、非常に多くの専門化された分野に分けられている。したがって材料科学の研究者でさえ、非常に限られた材料分野に関する知識を有していることが少なくない。そのため、材料情報システムには、材料に関する単なる特性情報だけでなく、ユーザがその情報を理解・利用するためのバックグランド知識も提供する必要がある。

（2）実験データ、計算データと理論知識の相互補完

実験、コンピュータ・シミュレーションおよび理論研究は、材料科学の三つの基本的な研究手法である。それらには、それぞれ特徴と制限があり、お互いに補完することも可能である。

例えば、理論知識は実験の指針であり、コンピュータ・シミュレーションのモデルとアルゴリズムの基本である。また実験データは原理、ルール等を抽出するための情報源であり、コンピュータ・シミュレーションの初期値としても利用される。さらにコンピュータ・シミュレーションは、実験が困難な条件の下でデータを取得することができ、実験データを補足する重要な手段となる。このように、実験データ、計算データと理論知識の総合利用は、最大限に材料知識を活用し、新たな知識を創出するためのポイントである。

（3）材料知識の発展性

新材料の開発や、理論の進展等に伴い、新しい材料データと知識は、常に創出されている。そのため、材料知識の管理システムは、拡張性を持ち、オープンなシステムではなければならない。

3.現在の材料知識管理システム

3.1 材料データの管理

現在、結晶構造データベース、合金データベース、高分子データベース等様々な材料データベースシステムが次々開発されてきており、飛躍的に増加しつつある材料の構造や特性等のデータの管理が行われている。それらシステムには、データの蓄積やインデックス、検索機能等しか有しない単純なデータベース型（図１）と、データベース機能に加えて、専用の解析・統計等データマイニングツールを持ち、データの統計等ができる拡張データベース型（図 2）がある。

Knowledge Management System for Materials Science
Materials Information Technology Station, National Institute for Materials Science

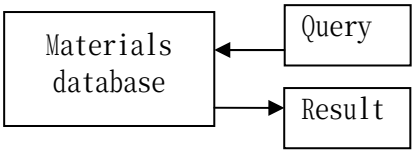

図１単純データベース型材料知識管理システム

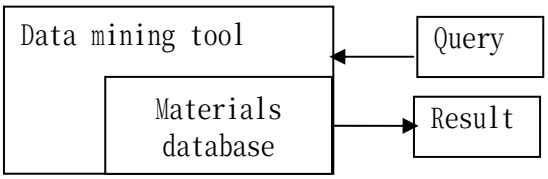

図２ 拡張データベース型材料知識管理システム

その一例として、日本科学技術振興団(JST)とMaterial Phases Data System (MPDS)が共同開発した合金材料データベースシステム Pauling File(http://lpfb.jst.go.jp/) を挙げる。Pauling File は、材料データ検索と材料設計支援ツールの 2 つの主要機能を有している。ユーザの指定した材料のＸ線回折データや、結晶構造、特性などを検索できるほか、データベースに保存されている全ての材料の構成、構造、特性等の統計結果をグラフや表等で表示することも可能である。

3.2 材料理論の管理

材料科学の概念や、原理、ルール等の理論知識を管理するシステムは、現在私たちの知る限り、未だ行われてはいない。

様々な材料シミュレーションと計算ソフトウェアは、材料科学の理論に基づいて、材料の構造や特性等のデータを計算するものであり、原理や方程式等大量な理論知識を含んでいて、その知識をコンピュータプログラムの方式で表現したものとも考えられる。しかし、知識管理の角度から見ると、これらシステムは、全てクローズな世界、言い換えれば知識の追加や検索・再利用が不可能であり、理論知識を管理・提供するための知識ベースとはなっていない。

4.現在の材料知識管理システムの問題点

現在の材料知識管理システムは、材料データしか管理しておらず、理論知識ベースがないため、情報の提供には、以下の問題点がある。

(1)知的な情報提供ができない。

提供できる情報は、材料データベースに保存されているものに限られているため、データベースにない材料データを理論知識に基づいて、自動推論により作成・提供することができない。例えば、データベースに材料の熱伝導率(κ)、比熱(c_p)、密度(ρ)のデータがあっても、熱拡散率($\lambda = \kappa / c_p \rho$)のデータがない場合には、熱拡散率の問合せに対して、何の情報も提供できない。

(2)提供する情報に関してのバックグランド知識サポートがない。

データベースに管理されている材料特性等に関連する知識の提供がないので、この分野の専門知識を持っていないユーザにとっては、理解しにくいことがある。

5.新しい材料知識管理システム

5.1 システム構成

上記の問題点を解決するために、データ管理と理論知識管理を一体化した新しい材料知識管理システム(図 3)を提案する。

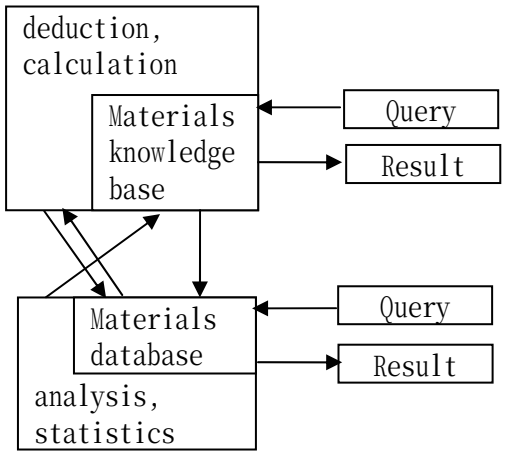

図３ データ管理と理論知識管理の一体化材料知識管理システム

図 3 のシステムでは、材料データと理論知識をそれぞれデータベースと知識ベースで管理する。知識ベースは、データベースの内容を正確に理解するためのバックグランドの知識を提供し、また、推論と計算機能により、データベースに含まれていないデータを算出することを可能にする。データベースは、計算に必要なパラメータや初期値等を提供し、さらに、データマイニング機能により、新しい原理やルールを発見することもできる。知識ベースとデータベースはそれぞれの検索インターフェースを持ち、ユーザの目的に合わせた最適な情報提供が可能である。

5.2XML での材料データの表現・管理

材料データとして表現される主要な情報内容を、図４に示す。

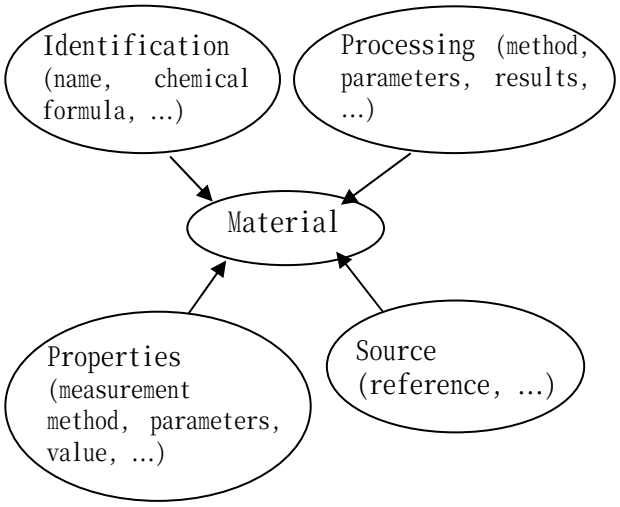

図４　材料データの情報内容

材料データの全ては、特定の材料に帰属している情報なので、XML の階層構造は、その所属関係を自然に表現できる。また、それぞれの材料に対して、必要とされる情報が違うので、データ項目もそれぞれ異なり、データ項目の追加・変更を行う場合が多く、データ構造の柔軟性が良い XML を用いなければ容易に対応できない。

材料特性データの XML 標準として、2001 年 3 月にアメリカの National Institute of Standards and Technology (NIST)により発表された MatML DTD[1]と、Europe と日本のニューマテリアルセンター (NMC)により発表された NMC's XML-DTD がある。我々はそれらを利用して、材料データを XML 形式で表現、管理するシステムを構築することが可能と考えている。

5.3 知識ベースにおける理論知識の表現

材料科学の理論は、概念の定義と概念間の関係を述べる原理や、ルール、方程式などから構成されるので、その表現は、材料科学概念とその相互関係のオントロジ表現である。

オントロジの表現に関して、たくさんの研究が行われていて、様々な技術が開発された。特に、よく知られているロジック規則の定義言語 The Knowledge Interchange Format (KIF)[2]や、近年に急速に発展してきた XML に基づいたオントロジ技術（RDF[3]、DAML+OIL[4]、OWL[5]など）などは、有力なオントロジ表現ツールとして注目されており、材料科学の知識表現に利用できる方法と考えられる。

次に、具体的な例を通して、RDF で材料知識のオントロジを表現する可能性を示す。

複合材料の熱伝導率については、次の知識がある[6]。“層状構造を持つ複合材料は、層状型複合材料と呼ぶ。Matrix-filler 構造を持つ複合材料は、分散型複合材料と呼ぶ。層状型複合材料の熱伝導率は、複合則で計算できる。分散型複合材料の熱伝導率は Hatta 方程式で計算できる。”

上記の複数の概念（複合材料、層状型複合材料、分散型複合材料、熱伝導率、複合則、Hatta 方程式）と、それらの相互関係を模式化して図 5 に示す。

図 5 に表わされる各要素は、XML 名前空間 **http://nims.go.jp/composite/** に所属すると仮定すると、図 5 に示した概念の定義と概念間の関係は、以下の RDF で表現できる。

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns ="http://nims.go.jp/composite/">
  <rdf:Description about="複合材料">
      <property>熱伝導率</property>
  </rdf:Description>
  <rdf:Description about="層状型複合材料">
      <superClass>複合材料</uperClass>
      <structure>層状構造</structure>
  </rdf:Description>
  <rdf:Description about="分散型複合材料">
      <superClass>複合材料</superClass>
      <structure>Matrix-filler 構造</structure>
  </rdf:Description>
  <rdf:Description about="熱伝導率">
    <calculationMethod>複合則</calculationMethod >
    <calculationMethod>Hatta 方程式
    </calculationMethod >
  </rdf:Description>
  <rdf:Description about="複合則">
     <applicableTo>層状型複合材料</applicableTo>
  </rdf:Description>
  <rdf:Description about=" Hatta 方程式">
     <applicableTo>分散型複合材料</applicableTo>
  </rdf:Description>
</rdf:RDF>
```

上記のように、RDF 技術を用いて、材料科学の理論知識を表現できる。しかしながら、その表現は、コンピュータの自動推論には、まだ十分ではない。自動推論のために、各概念および

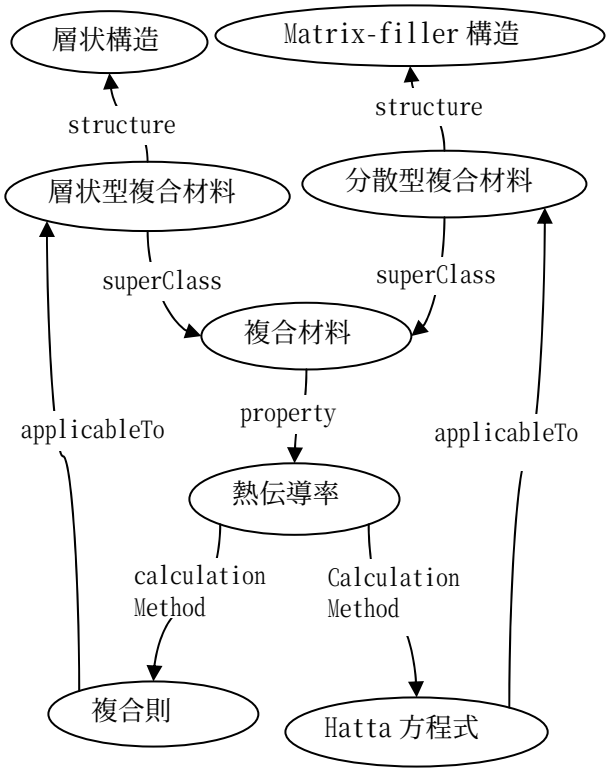

図５　複合材料の熱伝導率

概念間の関係をコンピュータの処理できるロジック表現も必要である。例えば、「A applicable to B」という関係に対して、「if B is true, then A is true」のロジック解釈が必要である。一方、ユーザにバックグラウンド知識を提供するために、理論知識を人間可読な形式で表現しなければならない。そのため、材料知識ベースは、理論知識をマルチ階層で表現・管理する必要がある。（図６）。

このように人間可読性とコンピュータ実行性を有した形式により知識ベースを構築すれば、バックグラウンド知識の提供および知的な情報提供が可能となるであろう。

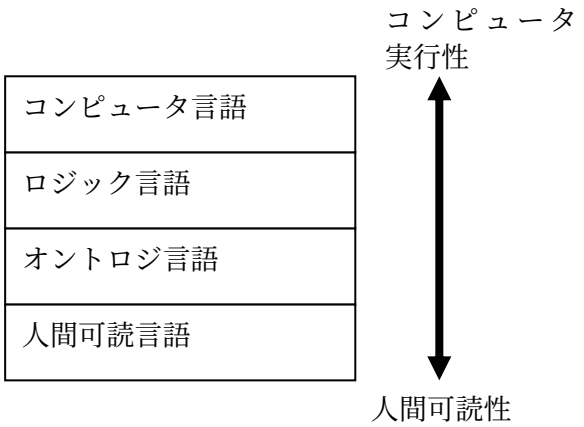

図６　理論知識の多階層表現

6.まとめ

従来の材料知識管理システムは、材料データのみに重点を置いたもので、知的な情報提供や、ユーザフレンドリーな知識サポート等の要求に対して、満足な回答を与えられなかった。本研究は、材料データと理論知識を統合した材料知識管理システムを提案し、材料データおよび理論知識の表現について検討した。

参考文献

[1] MatML DTD Version2.0, c National Institute of Standards and Technology, 2001. See ftp://www.ceramics.nist.gov/matml/MatMLv20.pdf

[2] M. R. Genesereth: Knowledge Interchange Format. See http://logic.stanford.edu/kif/dpans.html

[3] Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation 22 February 1999, c 1997,1998,1999 W3C (MIT, INRIA, Keio), 1999. See http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/

[4] DAML+OIL (December 2000), Joint US/EU ad hoc Agent Markup Language Committee, 2001. See http://www.daml.org/2000/12/daml+oil-index.html

[5] Web Ontology Language (OWL) Guide Version 1.0, W3C Working Draft 4 November 2002, c2002 W3C[R] (MIT, INRIA, Keio), 2002. See http://www.w3.org/TR/owl-guide/

[6] Y. Xu, J. Kinugawa and K. Yagi: Development of Thermal Conductivity Prediction System for Composites, Mater. Trans., To be published.

[7] J. F. Sowa: Knowledge Representation Logical, Philosophical, and Computational Foundations, published by Books/Cole, 2000.

[8] D. Fensel: Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce, published by Springer, 2001.