

無機材料分析手法の選択支援のための オントロジーの構築とシステムの提案

Building an ontology and proposing a system to support the selection of inorganic material analysis methods

砂田界渡¹ 來村徳信¹ 板倉修一² 笹本亮一² 細井慎²

Kaito Sunada¹, Yoshinobu Kitamura¹, Syuichi Itakura², Ryoichi Sasamoto², and Shizuka Hosoi²

¹ 立命館大学大学院情報理工学研究科

¹Ritusmeikan University

² 株式会社村田製作所

² Murata Manufacturing Co., Ltd.

Abstract: 本研究の目的は、無機材料の分析に際し、分析依頼書の内容から最適な分析手法を推薦するシステムに組み込むオントロジーを構築することである。分析依頼書に記載された分析目的を部分行為に分解し、その部分行為を達成する分析手法を出力するよう概念を定義する。加えて、その分析手法が用いる分析装置が適切かどうかを判断するために分析条件と結び付けて概念を定義する。これにより、分析目的が与えられたとき分析手法と分析装置の条件を結び付けて、最適な分析手法を推薦することを目指す。

1. はじめに

無機材料の製造において、物質の材質や内部構造の調査のために材料の分析が重要である。分析をするにあたって、共同研究先企業においては、分析依頼者は分析依頼書を作成して分析部門の技術者に依頼する。分析部門の技術者は、その分析依頼書をもとに、依頼者と分析相談会を行い、分析の依頼目的に応じた適切な分析手法を選択する必要がある。例えば、材料内の空洞(ボイド)の有無の調査が依頼目的であれば、「X線CT法」という分析手法を選ぶ必要がある。しかし、分析手法は共同研究先企業で用いているものだけでも約100種類あり、現状では特に経験の浅い技術者が数ある分析手法の中から最適な分析手法を選択することは、無機材料やX線・電子線といった使用媒体などの特性や依頼された調査目的、その分析手法によって明らかにできる性質などに基いて総合的に判断する必要があるため、困難な場合がある。

そこで、本共同研究全体の目標として、分析依頼書の内容から最適な分析手法を推薦するシステムの開発を目指し、本研究ではそのシステムに組み込むオントロジーを構築する。推薦システムはチャットボット方式を採用する。まず、ユーザは表示された分析目的行為の選択肢の中から分析の目的に合致す

るものを選択する。分析目的が確定した後、達成方式の選択肢が表示されるので、分析依頼に合致する方式を選択する。これにより達成方式と部分行為が同時に確定する。部分行為が確定した後、システムはオントロジーを参照し、分析装置の基本動作に選ばれた部分行為を持つ分析手法を自動的に複数抽出する。その後、分析装置の物理的な制約から作られる質問の選択肢に答えて分析手法を絞り込むことで、最適な分析手法を出力する。

分析の分類は観察・物性・組成・構造・化学構造・その他の6つとしているが、本研究で構築したオントロジーは、この分類のうち観察・組成分析を対象領域としている。

2. 注目する課題

本研究では、分析目的に応じた最適な分析手法を推薦するためのオントロジーを構築するにあたって、以下のような課題の解決を目指した。

第1の課題は、分析に関わる語彙定義や抽象度が不明確なことである。この課題はプログラムを複雑にし、メンテナンス性を低下させる原因となる。

第2の課題は、その分析手法が分析できることを記述すると記述量が増えスケールが大きくなってしまいうという問題である。例えば、走査電子顕微鏡法(SEM法(Scanning Electron Microscope))は、「表面の形

状を分析可能」,「ボイドの有無を分析可能」,「めっきの剥がれの確認可能」といったように複数の分析が可能であり,各分析手法の定義に分析可能事項を全て記述してしまうと記述量が非常に多くなってしまう.そのため,その分析手法が汎用的にできることのみを記述し,定義のスケールをなるべく大きくしないようにする必要がある.

第3の課題は,最適な分析手法と分析装置を推薦するにあたって,分析条件を考慮する必要があることである.例えば,走査電子顕微鏡という分析装置は試料を真空空間に置くため「試料は真空空間に置いても壊れないものである」という使用条件があり,「試料中に含まれるボイドの確認」といった分析目的に対して,解析試料が真空空間に置かれると壊れてしまう性質がある場合,解析試料が走査電子顕微鏡の使用条件を満たしていないので,走査電子顕微鏡を使用する分析手法は適切ではない,ということになる.従って,適切な分析手法を推薦するために,分析装置の使用条件を定義する必要がある.

3. 既存の分析手法の定義

3.1 RadLex

放射線分野において,北米放射線学会が放射線科領域に特化した用語集”RadLex”の作成を進めている[1].ここで定義されている用語は is-a 関係で整理された階層構造を為している.RadLex では,用語としての各分析手法は「Imaging Specialty(画像に関する専門技術)」-「Imaging Modality(画像診断法)」の下位概念で定義されており,主に観察を行う「Computed Tomography(CT 法)」や測定を行う「Spectroscopy(分光法)」などが定義されている.また,自然言語文によって概念の意味が定義されているものもあり,例えば,「CT 法」は自然言語文では「X 線透過とコンピュータアルゴリズムを使用して画像を再構成する断層撮影」と定義されている.

RadLex では統一された語の提供を行うが,機器の一般的な定義に加えて,「どの媒体を使用するのか」,「何を分析することができるのか」,「どのような分析結果が得られるのか」などの経験者が知り得る機器の特徴も重要である[2].そのため,本研究では,上記の特徴を含めて分析手法を体系的に定義する.

3.2 SNOMED-CT

SNOMED-CT は,医学用語を統一的に扱い表記を統一することを目的に,1999 年米国病理学会が定めた病理疾病分類 SNOMED をもとに,英国病院用の医療用語集 United Kingdom’s Clinical Terms Version 3(Read Code), Mesh(米国国立医学図書館科学用語

集)などを統合して作成された用語集である[3].概念としての各分析手法は「Procedure(処置)」-「Procedure by method(手法を用いた処置)」-「Evaluation procedure(評価に関する処置)」-「Imaging(画像診断)」の下位概念に定義されており,主に対象部位の画像検査を行う手法と特定の技術を利用する手法が定義されている.また,「Imaging」の下位に「Radiographic imaging procedure(放射線を用いた画像診断手法)」が定義されており,これの下位概念には,主に観察を行う「Computed Tomography(CT 法)」や測定を行う「X-ray photon absorptiometry (X 線光子吸光測定法)」が定義されている.

SNOMED-CT はターミノロジーとして発展してきたため,オントロジーとして見た場合,様々な課題が指摘されている[3][4].一例を挙げると,目的と手段の関係が is-a 関係で記述されているため,混乱を招く記述がされている.例えば,上述の”Imaging(画像診断)” is-a “Evaluation procedure(評価処置)”の部分は,画像診断によって「評価」を目的として達成できるということを表していると考えられる.本研究では,推論システムに組み込むことを想定したオントロジーを構築するため,分析目的に関する行為(全体/目的行為)とその目的を達成する行為(部分/手段行為)を分けて,ロール概念という基盤概念を用いて,記述する.SNOMED-CT の上述の is-a 関係は,目的-手段関係を is-a で表現しており,表現が十分ではなく区別が付かないため,推論時に混乱を招く記述となってしまう.

また,中間概念が豊富ではなく,用語の平行な列挙に近い構造になっている.例えば,「Radiographic imaging procedure」の下位概念には,X 線を照射して透過像を得る「CT 法」と,膀胱内に造影剤を入れて X 線撮影を行う「Pneumocystogram(膀胱造影法)」が平行に列挙されており,放射線の種類による分類と,対象臓器の種類による分類が区別されていない.つまり,オントロジカルな概念体系になっておらず,知識の再利用がしづらい構造になっている.

そのため,本研究ではオントロジーエディタの「法造」[5]を使用して,ロール概念と is-a 関係を明示的に記述することに加え,概念を意味的に分類し,知識の再利用を視野に入れた体系的な概念構造を構築する.

3.3 Measurement Method Ontology

生物医学分野では,表現型測定データを統合する必要性が高まっており,その中で「どのように測定したのか」に関する用語を標準化した Measurement Method Ontology が開発されている[6].「in vivo

method(生体内法)」と「ex vivo method(生体外法)」という2つの主要な概念があり、用語としての各分析手法はどちらかの下位概念で定義されている。例えば、「computed tomography(CT法)」は「in vivo method」-「in vivo radiography(生体内放射線検査)」-「tomography(断層撮影)」の下位概念で定義されている。

Measurement Method Ontology は、多くの分析手法が網羅的に定義されているが、その分析手法が使用する媒体(放射線)や基本動作で分類されておらず、使用技術ごとにパラレルに定義されたターミノロジーとなっている。そのため、SNOMED-CTと同様、体系的な概念構造になっておらず、知識の再利用がしづらい構造になっている。そのため、本研究では、3.1節で述べたように、分析手法を「使用媒体(放射線)」、「基本動作」、「分析結果」などの情報を含めて体系的に定義する。

4. 分析目的に関するオントロジーの構築

本研究で構築したオントロジーは上位オントロジーYAMATO[7]に準拠しており、それを一部簡略したものになっている。その理由は、行為の分解関係を「達成方式」という概念を用いて明示的に記述する方法を提供していることと、情報の表現物を「表現内容」と「表現媒体」に明確に区別して表現できる[8]からである。

オントロジーの知識源として、各分析手法の詳細が紹介された書籍[9]、共同研究先から提供された分析手法一覧ファイル、過去の分析事例に関するファイルを用いた。

本論文では分析に関する3つの基本行為(「観察する」と「測定する」と「加工する」)、その下位概念と分析目的との関係性、装置の機能や制約に関わる基本的概念を定義するための枠組みに焦点を当てて説明する。

本章では分析目的に関するオントロジーについて、分析行為と分析依頼目的の2つの概念を述べる。

4.1 分析行為

分析行為は、これ以上部分行為に分解できない分析に関する基本的な行為を指し、分析行為を「観察する」と「測定する」と「加工する」の3つに大別した。「観察する」を例に挙げると、図4.1に示すように、「観察する」の下位概念に「内部を観察する」を定義した。

まず、「分析行為」のスロットに、何を分析するかという入力物を表す「対象入力物」スロットを定義

した。「観察する」行為の対象入力物には、実際に観察する解析試料などが入るので、「分析行為」の下位クラスである「観察する」の定義において、「対象入力物」スロットのクラス制約を「もの」と記述した。

次に、「分析行為」のスロットに、その分析行為を行うことで何を得られるかを表す「対象出力物」スロットを定義した。「観察する」行為を行うと、試料の形状画像が観察結果として出力されるので、それを表現する「形態の表現可視化結果」をクラス制約に記述し、そのサブスロットに「画像」をクラス制約とした「形態」スロットと「観察結果」をクラス制約とした「分析結果」スロットを記述した。「内部を観察する」行為を行うと、観察した試料の内部形態を表す透過像が出力されるので、「対象出力物」スロットのクラス制約に「透過像」を記述し、そのサブスロットに「分析結果」スロットとしてクラス制約に「内部形態」を記述した。

さらに、「分析行為」のスロットに、「対象入力物」の試料のどの部分を分析するかを表す「測定位置」スロットを定義した。測定位置は「表面」、「裏面」、「内部」の3つがあり、例えば「内部を観察する」行為の定義では測定位置を「内部」と記述した。

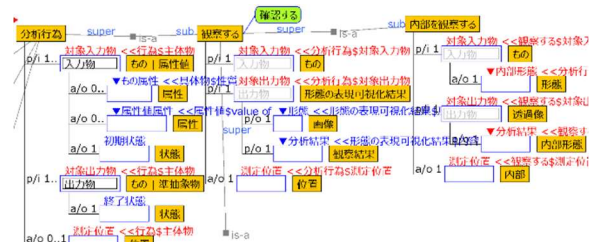


図 4.1 「分析行為」とその下位概念の定義

4.2 分析目的行為

分析目的行為は、分析依頼書に記述されている分析依頼目的に関する行為を指し、図4.2の「達成方式」スロットとして定義されているように、1つ以上の達成方式[10]にもとづいて、「分析行為 | 分析目的行為」をクラス制約とする複数の部分行為によって達成される行為である(「|」はOR関係を表す)。分析目的行為は「ものの形状を観察する」と「ものの属性を測定する」の2つに大別でき、例えば、「ものの形状を観察する」の下位概念に「物質を観察する」-「ボイドを確認する」を定義した。

ボイドとは、物質内の粒子と粒子の隙間のことをいい、「ボイドを確認する」は、ボイドが物質内に含まれているかどうかを確認することである。図4.2に示すように、「ボイドを確認する」という全体行為を達成する方式を2つ定義し、1つは「内部を観察する」ことで達成する「内部直接観察方式」、もう1つ

は「断面を出し」、その断面を「表面・裏面を観察する」ことで達成する「断面出し方式」を記述した。このように定義することで、1つの「ボイドを確認する」という分析目的行為を達成する方式が複数あり、それが加工行為である「断面を出す」や観察行為である「表面・裏面を観察する」のような部分行為をも含んでいるということを表現でき、分析目的行為(全体行為)と分析行為(部分行為)の関係性を明示的に記述することができた。



図 4.2 「分析目的行為」とその下位概念の定義

5. 分析手法に関するオントロジーの構築

本章では分析手法に関するオントロジーについて、分析装置と分析手法の2つの概念を述べる。

5.1 分析装置

図 5.1 に定義を示すように、「分析装置」は、主に、「分析行為」をクラス制約とする「機能」と、「放射線 | 化学物質」をクラス制約とする「入射物」「検出物」をスロットに持つ。分析装置は入射物に X 線を利用する「X 線利用装置」と電子線を利用する「電子線利用装置」の2つに大別でき、例えば図 5.2 に示すように、「X 線利用装置」の下位概念に「内部分析装置」-「X 線 CT 装置」を定義した。

X 線 CT 装置は、試料の内部を観察する機能を持つ装置であるので、図 5.2 に示すように「機能」スロットに「内部を観察する」を記述した。また、ボイド・固体物単体・部品を解析できる装置なので、「解析対象」スロットに「ボイド | 固体物単体 | 部品」を記述した。しかし、固体物単体であればどのようなものでも解析することができるわけではなく、X 線 CT 装置には分析条件があり、分析条件を満たした特定の性質を持つものしか分析することができない。このことは X 線 CT 装置のみならず、各分析装置にそれぞれ特有の分析条件がある。これを表現するために、図 5.2 に示すように「X 線入射装置」の「解析対象」スロットのサブスロットに「分析条件」スロットを記述し、第 6 章で述べる「解析対象物関連定性的属性」をクラス制約に記述した。例えば、X 線 CT 装置は分析条件として、解析対象物が密度不均一性を持つこと、装置の測定空間に入れた際に何らか

の影響で変色しないことが求められる。それらをそれぞれ「密度不均一性」、「非変色性」と定義し、「分析条件」スロットに、クラスそのものがスロットに入ることを表す「#オペレータ」を用いて記述した。この記述は 2 章で述べた 3 つ目の課題の解決に対応している。

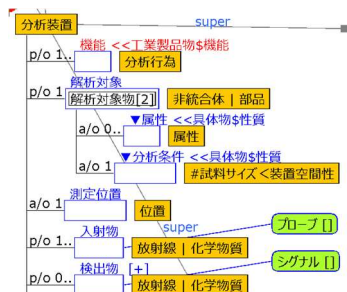


図 5.1 「分析装置」の定義

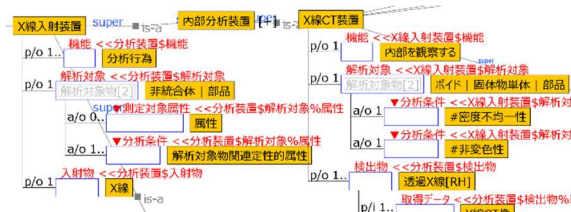


図 5.2 「X 線入射装置」とその下位概念の定義

5.2 分析手法

「分析手法」は、図 5.3 に定義を示すように、主に、「分析行為」をクラス制約とする「分析装置を用いて直接的に行える基本動作」、「分析装置」をクラス制約とする「用いる分析装置」、「放射線 | 化学物質 | 素粒子 | 熱プラズマ」をクラス制約とする「入射物」「検出物」をスロットに持つ。

分析手法も分析装置と同様に、入射物に X 線を利用する「X 線入射手法」と電子線を利用する「電子線入射手法」の2つに大別できる。例えば、「X 線入射手法」は図 5.4 に示すように、「入射物」スロットに「X 線」をクラス制約として記述した。また、X 線を利用した手法は物質を非破壊分析が可能なので、それを表現するために「分析特徴」スロットのクラス制約に「#非破壊性」を記述した。

「X 線入射手法」の下位概念に「内部分析手法」-「X 線 CT 法」を定義しており、これらの概念は上位概念のスロットで定義した「#非破壊性」を継承している。

X 線 CT 法は、汎用的に行える分析内容が、試料の内部を観察することなので、「分析装置を用いて直接的に行える基本動作」スロットに「内部を観察する」を記述した。この記述方法は 2 章で述べた 2 つ目の課題の解決に対応している。X 線 CT 法は「ボイドを確認する」、「物質内部の損傷を確認する」など様々な用途で 사용되는手法であるが、これらを

「基本動作」スロットに全て記述してしまうとスロットの定義量が爆発的に増えてしまう。そのため、汎用的に行えることを記述することで記述量を抑えられ、さらに7章で述べるオントロジーの活用法で、分析目的行為の部分行為を達成できる分析行為をスロットに持つ分析手法を適切な手法として候補に出力するというように処理することで、分析手法の定義のスケールを抑えつつ推論処理を可能とするオントロジーを構築することができた。

「X線CT装置」を用いて行う分析手法が「X線CT法」のように、本研究で対象となっている分析手法は全て「分析装置名=分析手法名」となっており、各分析手法はそれぞれ該当する分析装置を用いて分析する。そのため、各「分析手法」の「用いる分析装置」をロール名としたスロットのクラス制約に具体的な「分析装置」を記述した。また、装置の機能と分析手法が行う基本的な動作は一致する。例えば、X線CT法で分析できる内容は「内部を観察する」ことであるが、それはX線CT装置の機能が行うことである。つまり、X線CT法はX線CT装置を用いて直接的に行える基本動作として「内部を観察する」ことができると言える。このことを、「分析装置を用いて直接的に行える基本動作」と、用いる分析装置のスロットの「機能」を「sameAs」リンクで繋ぐことで表現した。

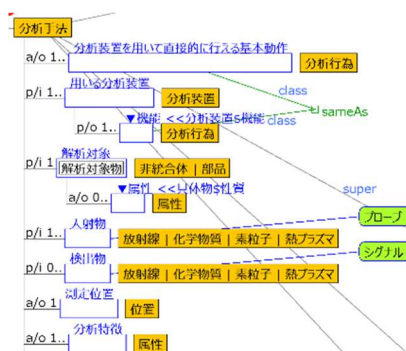


図 5.3 「分析手法」の定義



図 5.4 「X線入射手法」とその下位概念の定義

5.3 オントロジーの活用例

本節では、ここまでで述べたオントロジーをどのように使うことができるか具体例を使って述べる。「固体試料の中にボイドが含まれている可能性があり、ボイドの有無を調査したい。」という分析依頼に

対して、以下のようにオントロジーを使うことで、最適な分析手法を推薦することができる。

まず、分析目的から部分行為を抽出する。今回の例では「ボイドの有無を調査したい」が分析目的なので、まず、図 5.5(左)に示す「ボイドを確認する」という分析目的行為の定義が参照され、その達成方式・部分行為として、「内部直接観察方式」スロットの部分行為である「内部を観察する」と、「断面出し方式」スロットの部分行為である「断面を出す」・「表面・裏面を観察する」が抽出される。

次に、抽出した部分行為を基本動作に持つ分析手法が選択される。図 5.5(右)に示すように、基本動作が「内部を観察する」である「X線CT法」、「表面・裏面を観察する」である「走査電子顕微鏡法」と「電解放射-走査電子顕微鏡法」の3つが選択される。



図 5.5 オントロジー定義の参照と適切な手法の選定

ここから分析手法を絞り込む際、例えば「非破壊で分析する必要がある」という条件があったとする。この場合、「断面出し方式」の部分行為である「断面を出す」が図 5.6(右)に示すように「破壊性」のある行為なので、「断面出し方式」は破壊性のある分析行為を伴う。従って、最初の分析目的行為の達成方式とその部分行為を選択するステップで、「断面出し方式」から辿って選択された「走査電子顕微鏡法」と「電解放射-走査電子顕微鏡法」は適切な分析手法ではなくなるため、候補から除外される。一方で、「内部直接観察方式」から辿って選択された「X線CT法」は図 5.6(左)に示すように、「非破壊性」という分析特徴を継承しており、これは条件を満たしているため候補から除外されない。以上より、例に示した分析依頼に対する最適な分析手法は「X線CT法」となる。



図 5.6 「非破壊性」「破壊性」の定義

6. 分析条件に関するオントロジーの構築

分析条件は、解析対象物の定性的な属性で記述され、図 6.1 に示すように、「解析状況依存属性」と「解析対象物自身が持つ属性」の 2 つに大別した。

6.1 解析状況依存属性

解析状況依存属性とは、分析装置の測定空間内の環境に依存する条件である。例として、図 6.3 に示す「耐真空性」が挙げられる。これは、試料を解析する際、装置内の測定空間を真空にするため、試料は真空空間に置いても壊れないことが求められるという条件である。このように、装置の測定空間の環境が試料に影響を及ぼすような条件を「解析状況依存属性」の下位概念に定義した。

「耐真空属性」には次の 3 つを記述した。1 つ目は図 6.2 に示すように、解析対象物には内圧という圧力量を物質内に保持していることを記述した。2 つ目の記述として上位概念の「真空依存属性」のスロット情報を全て継承した。その際、物質の内圧の圧力量と装置空間が真空で圧力量が 0 であることを larger-than リンクで繋いだ。3 つ目に図 6.3 に示すように、力学運動の原理を記述した。ここでは、物質の内圧量が装置空間の圧力量よりも大きいことから、解析対象物の試料内部から装置空間に向かって力が働くことを force_from-to リンクを用いて記述した。これにより、物質内部で力がどのように働いているのかを原理レベルで定義することができ、次に述べる耐真空性の定義に際して、物質が真真空に耐えられることの厳密な力学運動を表現した。

「耐真空属性」の下位概念に「耐真空性」を定義した。「耐真空性」に、物質内部からの圧力量よりも物質の硬度量が大きいことで試料は壊れない、ことを表現するために、「硬度量」スロットと「圧力量」スロットを記述し、larger-than リンクを用いて、「硬度量」に larger、「圧力量」に smaller と繋いだ。

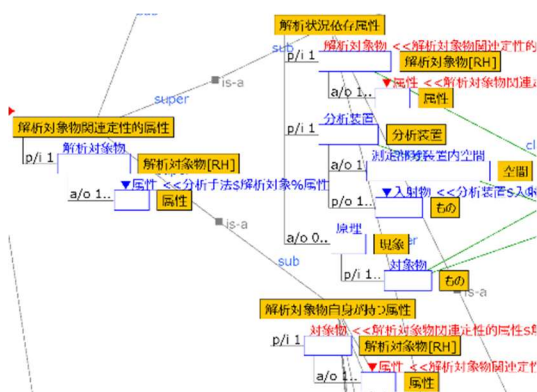


図 6.1 分析条件に関する 2 つの定性的属性

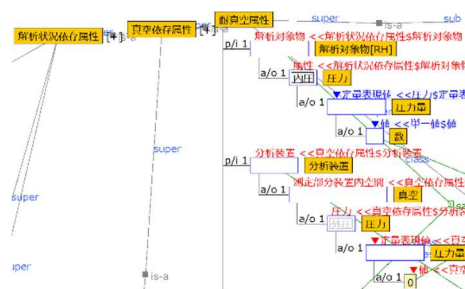


図 6.2 「耐真空属性」の定義(1)



図 6.3 「耐真空属性(2)」 「耐真空性」の定義

6.2 解析対象物自身が持つ属性

解析対象物自身が持つ属性とは、分析装置の測定空間内の環境に依存せず、試料そのものが持つ属性に依存する条件である。例として、図 6.4 に示す「耐熱性」が挙げられる。耐熱性は物質自身が持つ属性であり、耐熱温度の数値によって測定に影響が出るのは、試料そのものの属性に依存していると捉えられる。このように、装置の測定空間の環境に依存せず、試料自身が持つ属性で測定に影響が出る条件を「解析対象物自身が持つ属性」の下位概念に定義した。

「耐熱性」には、解析対象物自身が持つ耐熱温度に関する属性であることを表現するために、クラス制約を「耐熱温度」とした「属性」スロットを記述した。「耐熱温度量」サブスロットのサブスロットに具体的な温度量を表す「number」スロットと単位スロットを記述した。

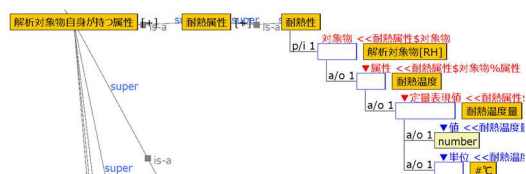


図 6.4 「耐熱属性」とその下位概念の定義

7. 推薦システムの実装と評価

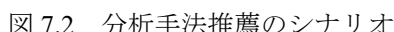
本章では、共同研究先で実装された推薦システムの説明とその評価について述べる。

7.1 推薦システムの説明

推薦システムの形式として、テキスト対話システ

オントロジーを利用したシステムの全体像を図 7.1 に示す。ユーザはチャットボットと対話を行う。どのような対話がなされ、対話でオントロジーのどこを参照したのかといったインタラクションは制御プログラムが管理する。制御プログラムはチャットボットとオントロジーの間に立ち、それぞれと通信をしながら、ユーザとの対話に基づいて上位概念から具体的な下位概念に辿ったり、対象とする装置の条件などを取得して必要な選択肢をチャットボットに表示を行うなどの制御を行っている。

ユーザが分析手法推薦プロセスで使うシナリオを図 7.2 に示す。このシナリオはオントロジーの構造に基づいて定義されている。ユーザは分析依頼内容に基づき、シナリオに沿ってリレーションを辿り、分析目的を確定する。そしてその分析目的に適合した分析手法と分析装置を絞り込むための条件を提示してもらう。さらにそこから絞り込みを通して提案の手法を選定する。



システムの利用手順を説明する。「エッチング工程でパターン作成を行ったときに残渣が出ているのが

次に検出感度の質問が表示される。候補の装置の属性から「0.1%～」、「0.5%～1%」と「不明」が表示

される。分析依頼書には元素の量についての記載はなく検出感度の条件は読み取れないため、「不明」を選択する。

絞り込みの質問に対し依頼書から読み取れた内容から回答が行われ、絞り込み条件がすべて揃ったため、絞り込みの質問は以上となる。質問の回答に基づいて絞り込みが行われ、システムは分析手法に「エネルギー分散型 X 線分析法 (SEM-EDX 法)」と「波長分散型 X 線分析法 (WDX 法)」を表示し、ユーザに 2 つの分析手法を提案した。この 2 つの方法を提案するのが正しい動作であり、以上で分析手法の絞り込みまで完了した。

7.3 オントロジーとシステムの評価

本節では、評価手順と評価結果について述べる。評価に際して、データベースから分析依頼書を抽出し、システムを使って分析手法と装置の提案をさせ、報告書と突き合わせて提案されたものが適切かどうかを判断するという方法とした。分析依頼書を抽出するにあたって、次の条件とした。まず、オントロジー辞書で定義された範囲内の文書に限定した。例えば、オントロジーで定義していない熱物性の測定や構造解析に関する文書は除外した。また装置まで定義をしていない分析手法が分析に使われるものも除外した。次に、分析部署が行う業務の範囲内とし、例えば組成分析でも外部委託を行ったようなものは評価対象から除外した。また、分析が複数の分析手法(定義された顕微鏡観察と組成分析に関する手法)を採用するケースは評価対象に含めた。最後に、別のサンプルのリピート依頼はサンプルに含まれないようにした。この条件で抽出された文書からランダムに抽出し、システムを用いて評価する。

評価手順は次の通りである。初めに、依頼書から依頼者が入力した項目を読み込む。このとき、文中に分析手法や装置の指定があっても隠して読む。試料の説明なども読み、試料形態であれば有機物・無機物の切り分け、固体や部品、液体などの分類を知るようにし、記述された内容からは無理な推測は行わないこととする。

次に 8.2 節で述べたシステムの利用手順に沿ってシステムを操作し、表 7.3 に示す評価項目表に基づいて採点する。基本は、システムで操作し読み取った内容をもとにシナリオをトレースする。ここで、オントロジー辞書の評価に注目をするためにプログラムの実装の不備やバグがあった場合は辞書を読み解いて補完するものとする。分析依頼書の依頼側入力から読み取れないものは選択肢から「不明」を選択する。先に進めなくなった(適切な選択肢がない)

ときはその次のプロセスの開始点に飛んで以降の評価を行う。評価は問題があった場合にシステムとオントロジーのどちらに起因するものかを明確にするため、それぞれについて実施する。例えば、オントロジーに問題がなくプログラムの不具合により操作等に問題がある場合は、システム側の問題として評価をする。

表 7.3 評価項目表

分類	項目
前工程 候補の抽出	1.1 必要な分岐・選択肢があるか?
	1.2 選択肢の意味が明確か? (戻ってやり直しが起きなかったか?)
後工程 候補から提案への 絞り込み	2.1 必要な絞り込み項目があるか?
	2.2 適切に絞り込めたか (明らかに不適切なものが提案に出ていないか?)
	2.3 絞り込み条件に不明で言えた後で絞り込みは問題が出ないか?
	2.4 手法が適切な精度などオーバーサペックになっていないか?

点数の付け方は次の通りである。評価の点数は各項目 10 点満点とする。ここから、上記手順で操作を行い、その際に発生した不適切な項目数の割合を引いた得点とする。例えば絞り込み項目 5 つが必要だが 1 つ定義が抜けていた場合は、4/5 が適切であったので点数は 10 点満点中 8 点とする。

評価結果を以下に述べる。まず、母数となる分析依頼書の件数、上記の抽出条件で取り出した評価対象の文書件数 (M)、サンプル数サイズ (N) を表 7.4 に示す。サンプリングは実際に適用された分析手法で比例層別サンプリングしてからランダムに抽出した。時間の関係もあり、実際に評価したサンプルサイズ (N) は M 数の 1% である。

表 7.4 母数・評価対象及びサンプルサイズ数

集合の説明	件数 (件)
データベース登録の分析依頼書件数 (母数)	110,028
オントロジー定義の範囲に含まれる評価対象文書件数	16,603
サンプル数	163

次に、評価結果を表 7.5 に示す。オントロジーの評価項目ではどの項目も 9.9 点以上と高い得点であり、システムの評価でも 9.5 点以上となっている。システムにはオントロジーがシナリオとして組み込まれているが、双方の点数には大きな乖離がないことから、シナリオの妥当性とオントロジーのシステムへの実装が適切であることを示している。またすべての評価項目で平均得点は 9 点以上であり、候補の抽出 (提案) や絞り込みで、過去の分析依頼に対して概ね適切に処理ができているといえる。

表 7.5 評価結果

	分類	項目	点数	
			オントロジー 評価	システム 評価
評価項目	前工程 候補の抽出	1.1 必要な分岐・選択肢があるか?	9.99	9.95
		1.2 選択肢の意味が明確か? (戻ってやり直しが起きなかったか?)	9.98	9.75
	後工程 候補から提案への 絞り込み	2.1 必要な絞り込み項目があるか?	9.90	9.51
		2.2 適切に絞り込めたか (明らかに不適切なものが提案に出ていないか?)	9.97	9.88
		2.3 絞り込み条件に不明で言えた後で絞り込みは問題が出ないか?	10.00	10.00
		2.4 手法が適切な精度などオーバーサペックになっていないか?	10.00	10.00

7.4 評価結果の考察

システムを使うにあたって、重篤性の高い順で述べる。まず、候補の提案結果に問題あるケースについて、問題点を2つ述べる。

問題点の1つ目として、分析目的で行き止まりになることがあった。前工程の分析目的行為の絞り込みで、下位概念に落とし込んでいくプロセスにおいて、分析目的行為の「素子を観察する」から部分行為を基本動作に持つ分析手法に辿ることができなかった。

問題点の2つ目として、未登録の絞り込み項目があった。例えば、WDX法では一度に認識できる元素は5つであるが、これがオントロジーに登録できていなかった。このためSEM-EDXとWDXのどちらを選択するか判断ができなかった。

次に、本評価実験の前提に照らして提案結果には合理性があるが、操作時に戸惑ったケースを4つ述べる。

1つ目は、誤った選択肢を選び元に戻ってのやり直しが発生したケースである。システムがユーザに提示する選択肢はオントロジーで定義した語彙を使うことが多いため、適切な言葉が無く分かりづらい表記が出てきて、誤ったものを選んでしまうことがあった。例えば、部品・素子・物質という語彙が出てくるが、素子に異物が付着した場合などは、部品なのか、素子なのか迷うところである。

2つ目は、複数の分析手法を組み合わせるケースである。例えば、表面観察と、その結果でいて組成を調べる場合はそれぞれに対してチャットボットを操作する。表面観察の結果が出るまで組成分析の条件の一部が確定しないため分析手法が絞れないことがあった。

3つ目は、一次スクリーニングのケースである。SEM-EDX法は組成分析の一次スクリーニングに使うこともあり、そのような分析手法であるという理解が必要であるが、オントロジーでは定義していなかった。

4つ目は、論理的な判断と異なるケースである。分析部門は、装置の空き具合やコスト、測定に掛かる時間などを考慮して適切な分析手法・分析装置を提案するため、状況によっては必ずしもロジカルな判断が正しい選択にはならないことがある。

最後に、提案結果に問題は無いが、システムの改善が望まれる点を2つ述べる。

1つ目に、条件による絞り込みの順番が不自然に感じられることがあった。辞書には順番の定義はなく、提示された分析手法の組み合わせによっても順

番は違うので、その定義は一律には決められないが、不自然に感じる場所があった。

2つ目に、部品・素子・物質などの区別が難しい場合にエイリアスを使って例示を登録し、それを表示する方法を取った方が良いケースがあった。

上述の問題点はオントロジー辞書の構造を変更することなく概念の修正と登録追加で解決できることから、オントロジーの構築とプログラムへの実装はうまく行ったといえる。

7.5 今後の展望

取得したサンプルデータの平均値は9.910で、下側95%信頼区間を計算すると9.788であり、評価対象Mでも高い平均値であると考えられる。サンプルデータは得点10.00に偏りが大きい傾向がありバイアスの影響の可能性が否定できないが、評価対象(M)の推定得点平均値から考えると残りもこのオントロジーの定義で適切な処理ができると予想できる。

さらに全体に対しての展望であるが、ここで評価対象としたのは母数110,028のうちの16,603であり全体の15%のみである。しかし分析手法提案シナリオや装置の機能や原理などの複雑な構造設計は残りの85%にも適用可能なものであり、このことから全体のかかなりの部分のオントロジーの構造設計は完了している。今後のオントロジー拡張では構造設計が少なく、登録作業が主となるため、15%に対する85%の比率に比べて大幅に少ない工数でオントロジー構築ができると見込まれる。業務全体をカバーするのに個々の変更事項はあろうが、シナリオ自体は問題なさそうであり、オントロジーで修正・追加等を進めていくことで実務に耐えるシステムになると考える。

8. 関連研究と考察

本章では、本研究と他分野で構築された推薦オントロジーを比較し、特徴を考察する。

本研究は、分析目的から最適な分析手法を推薦するためのオントロジー構築において、次の2つの特徴がある。1つ目は、分析目的行為を部分行為に分解し、その部分行為を基本動作に持つ分析手法を出力するように構築したことである。「部分行為」ルールと「基本動作」ルールに入るクラス制約を分析行為としたことで、2章で述べた第2の課題であるスケール爆発問題の解決につながり、同時に3.1節で述べた従来オントロジーの欠点である分析特徴の未記述問題や3.2節で述べたis-a関係のみの定義による目的—手段関係の不明確問題を解消した。2つ目の特徴は、分析条件を考慮して分析手法と分析装置を

出力するように構築したことである。分析条件オントロジーの構築と分析装置の分析条件スロットの記述により、2章で述べた第3の課題である分析条件の考慮問題を解決し、分析目的が与えられたとき分析条件を結び付けて「最適」な分析手法と分析装置を推論できることを示した。この2つの特徴を取り入れたオントロジーの構築により、2章で述べた第1の課題である分析に関わる語彙が定義や抽象度が不明確である問題を解消し、推薦システムを開発することができた。

他分野の推薦オントロジーとして、加藤ら[11]は、日本語 WordNet, LinkedGeodata, DBpedia といった大規模なデータベースを用いた知識推論を提案している。この研究では、自動車搭乗者が持つ曖昧な個人的要求から具体的な目的地を推薦するためのオントロジーを構築している。例えば、「ライオンが見たい」という曖昧な要求に対して、行動対象オントロジー・サービスオントロジー・施設オントロジーを用いて、最終的に「動物園」を推薦している。具体的には、まず行動対象オントロジーと日本語 WordNet を用いて、「行動=見る」、「行動対象=ライオン」を同定する。次に、サービスオントロジーと施設オントロジーを用いて、「行動=見る」、「行動対象=動物」を満たす施設クラス(動物園)を決定する。最後に、LinkedGeoData, DBpedia を利用して、現在地から近い施設のインスタンスとその情報を出力する。これにより、オントロジーで適切なクラスを推薦するだけでなく、大規模なデータベースを活用することで適切なインスタンスとその情報まで推薦できるため、ユーザに納得感を与える。

本研究は全体行為を分解し、その部分行為を満たす分析手法を適切な手法として出力するものであり、全体行為を直接的に満たす手法を出力するのではなく、分析目的を達成する部分行為に着目して分析手法を推薦するアプローチをとっている。そのため、多様な分析目的(全体行為)の実現に対して部分行為まで掘り下げて推論でき、ユーザは出力された手法がどのようなプロセスを経て導き出されたかを理解しやすくなり、ユーザに更なる納得感を与えることができる。加えて、手法・装置の持つ分析条件を考慮して推論でき、より複雑な推論を可能としている。

9. おわりに

本研究では、最適な分析手法を推薦するための概念を定義し、オントロジー・推薦システムの評価を行った。

今後の課題として、分析行為の「加工する」の下位概念の拡充が挙げられる。現状では、「観察する」と「測定する」の下位概念は充実しているが、「加工

する」の下位概念は「断面を出す」行為しか定義していない。分析依頼には前処理や後処理で加工行為を伴う場合があるので、今後はその加工行為を「加工する」の下位概念に定義したり、その加工行為を定義済みの分析目的行為の部分行為に追加したりして、より多くの分析依頼に対応できるようにする必要がある。

参考文献

- [1] Langloz CP.: RadLex: a new method for indexing online educational materials. Proceeding of Radiographics, Vol.26, No.6, pp.1595-1597 (2006).
- [2] 福田晋久, 辻真太郎, 谷川原綾子, 西本尚樹, 小笠原克彦: 放射線技術分野への応用に向けたオントロジーの分類, 日本放射線技術学科雑誌, 71 巻, 6 号, pp.505-511 (2015).
- [3] 藤田伸輔, 今井健: SNOMED-CT と ICD-11 に見る医学・医療分野の Ready to Use Ontology, 人工知能学会誌, 25 巻, 4 号, pp.501-508, (2010).
- [4] Schulz, S., Suntisrivaraporn, B., Baader, F.: SNOMED CT's problem list: ontologists' and logicians' therapy suggestions, Stud, Health Technol. Inform., Vol. 129, Pt 1, pp.802-806 (2010).
- [5] 古崎晃司, 來村徳信, 池田満, 溝口理一郎:「ルール」および「関係」に関する基礎的考察に基づくオントロジー記述環境の開発, 人工知能学会論文誌, 17 巻 3 号, pp.196-208 (2002).
- [6] Shimoyama M, Nigam R, McIntosh LS, Nagarajan R, Rice T, Rao DC, Dwinell MR.: Three ontologies to define phenotype measurement data, PMID: 22654893; PMCID: PMC3361058 (2012).
- [7] Mizoguchi, R., Borgo, S., YAMATO: Yet-another more advanced top-level ontology, Applied Ontology, Vol.17, No.1, pp. 211-232 (2022).
- [8] 溝口理一郎: オントロジー工学, オーム社 (2005).
- [9] 一般財団法人材料科学技術振興財団, 分析手法カタログ Ver.7, 2017 年 11 月発行号
- [10] 來村徳信, 溝口理一郎: オントロジー工学に基づく機能的知識体系化の枠組み, 人工知能学会論文誌, 17 巻, 1 号, pp.61-72 (2002).
- [11] 加藤文彦, 小出誠二, 武田英明, 落合勇太, 上田建揮: オープンなデータベースを利用した行動計画提案に関する研究, 人工知能学会全国大会論文集, 名古屋, 1N2-OS-39a-2in1 (2017).