

オントロジー工学的立場から見たオントロジー利用研究の分類と考察^{1D3-3}

A Classification and Consideration of Utilization of Ontology based on Ontological Engineering

古崎 晃司^{*1}
Kouji Kozaki

來村 徳信^{*1}
Yoshinobu Kitamura

溝口 理一郎^{*1}
Riichiro Mizoguchi

^{*1} 大阪大学産業科学研究所
I.S.I.R., Osaka University

Recently, ontology is attracted as one of key technologies for knowledge engineering and semantic web. As the result, many systems using ontologies are developed. However it is difficult to understand roles of ontologies in those applications because ontologies are used in many different ways. The authors analyze utilization of ontologies and classify them based on ontological engineering. And this article discuss a trend of applied research of ontology.

1. はじめに

近年のオントロジー研究の広がりに伴い、オントロジーを利用したシステムの開発事例も多く見られるようになった。しかしその反面、オントロジーに関する意見は、肯定的なものから批判的なものまで様々な意見が聞かれる。この一因は、オントロジーの意味合いやオントロジーの利用形態が多様性に渡っていることにあると考えられる。ライトウェイト - ヘビーウェイトオントロジーという言葉はその象徴で、ライトウェイトな点だけを見て批判的な意見を持ったり、ヘビーウェイトな点に重要性を見いだしたり、各自が理解しているオントロジーの違いによって、意見が分かれるのは当然であろう。しかも、ヘビーウェイトなオントロジーを構築するコストを問題視し、ライトウェイトなオントロジーの利便性を好む立場もあり、これらのオントロジーを取り巻く状況を考えると、一般の利用者からはその実態がつかみきれないように思われる。

そこで本論文では、オントロジー工学に携わる研究者の立場からオントロジーの様々な利用法を考察し、特にオントロジーの応用研究事例を対象として、様々な立場の違いやその傾向を分析することを試みる。そして、その考察を通して、今後のオントロジーの応用研究の展望について議論する。

まず、様々なオントロジーの応用研究の分析を進めるにあたり、オントロジーの利用法を分類するための指標として、1) 利用するオントロジーの種類、2) オントロジーの利用形態という2つの観点について2章および3章で議論する。続く4章では、これらの2つの指標を組み合わせて、「どのようなオントロジーをどのような目的で利用しているか」という観点から、既存のオントロジーの応用研究を分類し、その結果を考察する。そして5章で本論文の総括を行う。

2. オントロジーの種類

オントロジーには、上位オントロジー、タスクオントロジー、ドメインオントロジーといった代表的な種類や、ゲノムオントロジー、プラントオントロジー、機能概念オントロジーなど、対象とする内容による分類が考えられる。しかし本論文では、上述のようなオントロジーの種類や内容に依らない共通の分類の観点として、

(1) オントロジーの意味的構成要素

(2) 哲学的な考察の深さ

という2つの観点からオントロジーを分類する。しばしば見られる、ライトウェイト(Light weight)オントロジー、ヘビーウェイト

(Heavy weight)オントロジーといった違いは、これらの分類軸を合わせて検討したものと考えられる。

2.1 オントロジーの意味的構成要素

ここでは、オントロジーが対象としている内容を構成している意味的要素の種類によって、オントロジーを分類する。ここで述べるオントロジーの構成要素は、本質的には、オントロジーの記述言語と独立して議論すべきである点に注意して頂きたい。それは、記述言語によって多少の限界はあるものの原則的にはどのような記述言語を用いても、構成要素の内容を記述することは、ほぼ可能だからである。ただし、オントロジー記述言語によっては、あらかじめオントロジーの内容を表すのに必要となる構成要素をあらかじめ言語として提供しているため、オントロジーを記述しやすい場合がある。その場合であっても、提供された構成要素を使用するか否かは記述者に任されるので、同じ記述言語で書かれたオントロジーが必ずしも、本節で述べる同じ種類のオントロジーに分類される訳ではない。

それでは以下に、オントロジーの意味的構成要素による分類を述べる。

(1) 統一された語彙集合 / 簡単なスキーマ

本来はオントロジーとは呼ばれないものであるため不適切ではあるが、様々な応用研究においてオントロジーが果たす役割と近い目的で利用されることが多いため、それらを合わせて比較するために、本論文では「統一された語彙集合」および「簡単なスキーマ」についても分類の一つに含める。

統一された語彙集合とは、知識を記述する共通語彙として用いるために用意された語彙の集合を指す。ブログのエントリーに用いるカテゴリや、論文投稿時に運営組織から提供されるキーワード一覧などがその典型例である。一方、簡単なスキーマとは、RSSやFOAFのように統一したデータ記述のために提供されるスキーマを表す。ここで重要な点は、単なる自由記述によるキーワードや、特定のデータベースのためだけに作られたスキーマとは異なり、知識の相互運用性を考慮して統一された形式として提供されていることにある。

なお、これらをより詳細化して行くことで、意味的に表す内容が以下に述べる(2)~(5)に近づくことは自明だが、その境界は恣意的なものでありここでは議論しない。

(2) 概念間の is-a 関係に基づく階層

対象世界に存在する概念を切り出した結果としての「概念」の集合と、それらの概念間の is-a (上位-下位) 関係を同定することによる階層化が、オントロジーにとって最も本質的な要素となる。

連絡先: 古崎晃司, 大阪大学産業科学研究所, 〒567-0047
大阪府茨木市美穂が丘 8-1, Tel:06-6879-8416, Fax:06-6879-2123, kozaki@ei.sanklen.osaka-u.ac.jp

概念階層という表面的な形式だけを見ると、シソーラスの上位語・下位語による階層や、分類階層(タクソミー: Taxonomy)と同等のものとして、is-a 階層のみから成るオントロジーをライトウェイトオントロジーと呼ぶ場合もある。Web 検索サイトのトピック階層はその典型例と言える。

しかし、「何を上位・下位概念とするか?」という考察はオントロジーにとって本質的で深い議論が含まれており、上位オントロジーのように 2.2 節で述べる哲学的な深い考察に基づくオントロジーが is-a 階層のみで表される場合もある。

(3) is-a 以外の関係を含む

is-a 関係に基づく概念の階層化に加え、それ以外の関係を用いることで各概念の定義をより明確に表すことができる。part-of(全体 - 部分)関係や、attribute-of(属性)関係が最もよく使われる関係である。シソーラスや分類階層とオントロジーの大きな違いの一つは、このような関係の記述があるか否かで捉えることもできる。RDF(S) が提供している rdfs:subClassOf と rdfs:Property は、is-a 関係とそれ以外の関係を記述するのに用いることができ、RDF(S) で記述されたオントロジーの多くはこの種類のオントロジーに分類されると思われる。

(4) 意味制約の公理的記述を含む

関係を用いた各概念の意味定義の記述に加え、意味制約の公理的記述を行うことで、インスタンスモデルにおいて適切な規約を与えることができる。ここで記述される意味制約には、推移律や反射律などの関係の性質、「同じ自転車の前輪と後輪は異なる車輪である」といったインスタンスの排他性に関する制約などが含まれる。意味制約を公理的に記述する表現能力はオントロジー記述言語の仕様に依るが、OWL を用いた場合は transitiveProperty, inverseOf といった関係の性質に関する制約や disjointWith, oneOf といった意味制約を用いることができる。オントロジーの本質である概念定義の大部分は、これらの意味制約を用いて表すことができる。

(5) その他の強い公理¹を含む

意味制約だけでは十分に表しきれない意味定義に関しては、ルール記述言語などを用いた強い公理とした記述が必要となる。KIF や SWRL などを用いたルール記述を伴うオントロジーがここに分類される。

2.2 哲学的な考察の深さ

前節で述べたオントロジーの意味的構成要素からみた分類は、オントロジーの構成要素を形式的に分類したものと位置づけることができる。それに対して、別の立場からオントロジーを分類すると、

- 1) 情報論的な利用効率を重視したオントロジー (ライトウェイトオントロジー)
 - 2) 哲学的な考察に基づき対象世界を適切に捉えることを重視したオントロジー (ヘビーウェイトオントロジー)
- という 2 種類に分けることができる。

前述のように、オントロジーの意味的構成要素の違いから、より複雑な意味記述を含むオントロジーをヘビーウェイトオントロジーと呼ばれる場合もあるが、ここで述べた 2 種類のオントロジーに対する立場の違いが、より本質的な意味でライトウェイト - ヘビーウェイトの違いを表していると思われる。しかし、その考察内

¹ 公理やルールという言葉が表す内容には、意味定義に用いる制約や、推論に用いる規則など多くの意味が含まれている。ここでは厳密な議論を避け、前述の 1) ~ 3) で表される内容以外を「強い公理」と呼ぶ。

容を適切に表すためには厳密な意味記述が必要とされるので、これらの立場の違いは、構築されたオントロジーの構成要素の違いにも反映される。よって、ライトウェイトかヘビーウェイトかといった議論は、双方の観点から総合して判断されるべきである。

3. オントロジーの利用形態

オントロジーをどのような目的で利用するかによって、使用されるオントロジーの種類や利用方法は大きく異なる。ここでは、オントロジーの主な利用形態について分類する²。2 章で述べたオントロジーの種類は、原則としてこれらすべての利用形態において用いられる可能性があり、同じ利用形態であっても用いられるオントロジーの種類によって扱える内容の範囲が異なる。

0) 辞書的な利用

自然言語処理やテキストマイニングなどを行う際の辞書としてオントロジーを用いる。典型的な応用事例としては、オントロジーを利用した検索システムや文献のクラスタリングツールなどがある。この目的においては、主に概念定義を与えるオントロジーのみが利用され、以下の 1) ~ 6) の目的と組合せて利用される場合を除きインスタンスモデルは用いられないことが多い。

1) 共通語彙としての利用

知識の相互性を高めるために、知識を記述する際に用いる共通語彙としてオントロジーを利用する。最も基本的なオントロジーの利用法であるため、以下で述べる 2) ~ 6) の利用目的においても共通語彙としての役割はオントロジー利用の大前提となる。それらの共通語彙は、検索のキーワードや、統計的手法を用いた処理の対象など、様々な用途で用いられるが、以下の 2) ~ 6) で述べる利用目的に特化したものについては、それぞれの分類に入れる。

具体的な利用法としては、ドキュメントへのタグや、データベースの値として用いられることが多い。Web 上で様々な情報にキーワードをタグ付けすることによって情報を管理するサービスが最近よく見られる。

2) インデックスとしての利用

適切な情報へアクセスのためのインデックスとしてオントロジーを利用する。この利用法は、オントロジーが提供する語彙体系をインデックス(見出し)として利用したもので、その語彙を用いて情報にアクセスする。体系の中における語彙の位置づけを明示的に利用する点が 1) の共通語彙としての利用と異なる。Web 検索サイトで用いられるディレクトリサービスがこの典型例である。また分散したドキュメントなどのコンテンツに付与したメタデータを、インデックスとして利用することもできる。このような利用形態の具体事例としては、自動リファレンスサービス[黒橋 01] やコンテンツ管理システム[Guarino 99, 古崎 06], Web サービス検索エンジン[川村 05]などがある。

3) データスキーマとしての利用

データベースなどに格納するデータの構造や値を規定する標準化されたデータスキーマとしてオントロジーを利用する。オントロジーが提供する概念階層がデータの値となるインスタンスを規定し、各概念定義がデータの構造を規定する。典型的な利用例としては、データベースにおけるスキーマや、ドキュメントに付与する書誌情報などといったメタデータの内容を規定するメタ

² 本分類は、[溝口 05] (5.4 節) や [Kitamura06] を参考にしつつ再構成したものである。

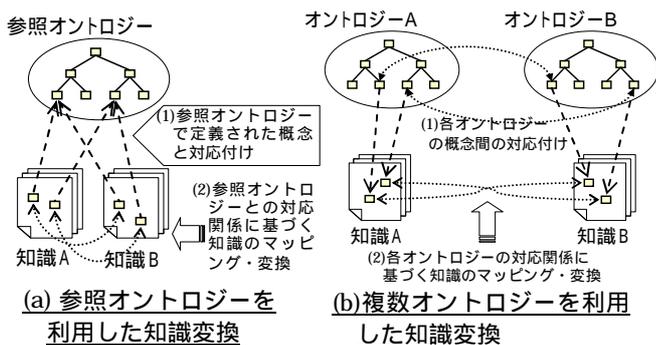


図 1. 知識共有の媒体としての利用

データスキーマとしての利用などがある。オントロジーをデータスキーマとして利用するメリットは、その相互運用性や意味構造を用いた知的処理(最も簡単例としては is-a 階層を用いた検索範囲の制御など)にあり、ヘビーウェイトなオントロジー(3.1 節参照)を利用することでより知的な処理が可能となる。

4) 知識共有の媒体としての利用

オントロジーおよびオントロジーに基づくインスタンスモデルを用いて、システム間、人間とシステム、あるいは人間同士の間で知識の相互運用性を高め知識共有を実現する。1)で述べた語彙の共有や、5)で述べる知識記述の規約としての利用も、知識共有の実現を目指したものと言える。しかし、ここでは複数システム間での知識の変換やマッピング、エージェント間の通信、コミュニケーション支援など、知識の相互運用性を中心的な目的としたオントロジーの利用を、このタイプに分類する。技術的には図 1(a)に示すように、複数の知識を同一のオントロジー(参照オントロジー)と対応付けることで変換やマッピングを行う手法や、図 1(b)のように複数のオントロジーを直接比較する手法などがある。その際、必要に応じてマッピングに用いる知識が参照オントロジーを利用して記述される。図 1 の知識 A,B をオントロジーとすると、複数の異なるオントロジーのマッピングや変換も、同様の枠組みでの実現が考えられる。

具体事例としては、データベーススキーマの変換、知識集約型工学のためのフレームワーク(KIEF)[吉岡 00]、離散型プロセス記述のためのオントロジー:PSL[PSL]、機能を中心とした設計知識の変換システム[小路 04]、異なる機能語彙体系のマッピング[大久保 06]などがある。

5) 知識モデルの規約としての利用

オントロジーが提供する概念定義に基づいて構築されたインスタンスモデルは、対象世界のオブジェクトを表した知識モデルとして用いることができる。すなわちオントロジーは、その知識モデルに対する規約を与えるメタモデルとして利用できる。これは 4)で述べたデータとデータスキーマの関係に相当するが、知識モデルには単なるデータに比べて柔軟な知識の意味記述が求められる。セマンティックウェブにおける RDF(S)や OWL を用いたオントロジーは、Web ドキュメントのメタデータとなる知識モデルに対する規約として利用されることが多い。また知識ベースシステムにおいては、メタデータとしてではなく知識モデルそのものを処理対象とするものも多くあり、それらのモデルに対する規約としてオントロジーを利用することは、知識の共有・再利用性に大きく貢献する。データと知識モデルの厳密な境界には恣意性があるが、一般的なデータベースのように知識の構造が固定的であるか、知識の記述に規約に反しない範囲での自由度が認められるかが、データと知識モデルの境界の一つの目安と考えられる。

具体事例としては、ゲノム分野におけるパスウェイの統一記述[高井 05]、e ラーニングにおけるオーサリングシステム[林 03]などがある。

6) 知識の体系化への利用

3.1 節で述べたオントロジーの意味的構成要素は、知識の体系化・組織化の核となる概念構造を提供する。それらを用いることで、対象世界の知識を構成している概念をオントロジーとして体系化し、その概念定義が与える規約に基づいて知識管理・運用を実現することができる。これは 1)~5)で述べてきたオントロジーの利用目的を総合的に用いたものと位置づけることができ、知識管理システムやコンテンツ管理システムといった統合的なシステムが、その典型的な応用例となり得る。

オントロジーを知識の体系化に用いた具体事例としては、エンジニアリングデータを規定する統一モデルである EPISTLE[岡田 04]プロジェクトや、筆者らが行ってきたオントロジーに基づく機能表現[来村 02a, 来村 02b]がある。

4. オントロジー応用例の分析

ここでは、2 章で述べたオントロジーの種類と 3 章で述べたオントロジーの利用形態を基に、近年のオントロジー応用研究の事例について分析を行う。

4.1 分析対象

分析の対象とする研究事例は、現在、J-Stage¹で公開されているデータベースから「オントロジー」をキーワードとした検索(検索対象は全文)を行い取得した文献(ジャーナル、学会発表予稿集など)、およびセマンティックウェブとオントロジー研究会(SIG-SWO)の Web サイト²で公開されている発表資料(第 1 回から第 14 回まで)を用いた。J-Stage から得た文献については、オントロジー研究に関するもの以外を除き、SIG-SWO のサイトから得たものと合わせて合計 223 件の文献を対象として分析を行った。

4.2 オントロジー応用研究事例の分析結果

前節で述べた文献を対象として、それぞれの研究について、(1)オントロジーの応用研究、(2)その他のオントロジー研究に分けた。さらに(1)に分類された 114 件について、2 章で述べた「どの種類のオントロジーを用いているか」と、3 章で述べた「オントロジーの利用形態」を用いて分類した。オントロジーの種類については、2.1 節の意味的構成要素の観点で分類した。表 1 に応用事例の分析結果を示す。これらの結果において、同一の研究内容が複数の文献で述べられている場合は、重複カウントを避け研究内容毎に 1 件とした。なお、対象とした文献の書誌情

表 1 オントロジー応用研究事例の分析結果

		オントロジーの種類(意味的構成要素)				合計
		0)語彙	1)is-a	2)関係	3)意味制約 4)強い公理	
論文 の 種 別	0)辞書的	2		3	1	6
	1)共通語彙	12	1	11	2	26
	2)インデックス		5	6		12
	3)データスキーマ	5		7		12
	4)知識共有		12	9	5	26
	5)モデル規約		2	22	7	31
	6)体系化				1	1
合計		19	20	58	16	114

¹ <http://www.jstage.jst.go.jp/>

² <http://www.jaist.ac.jp/ks/labs/kbs-lab/sig-swo/>

報を含めた結果の詳細は Web サイト¹で公開しているの、興味のある方はご覧下さい。

4.3 考察

ここでは前節で示した分析の結果について考察を行う。

利用されているオントロジーの種類

2)is-a 以外の関係を含むものが約半数を占め、4)強い公理を含むものは殆ど見られない。これにより、現状の応用研究では、強い公理がさほど必要にされていないことが伺える。

オントロジーの利用形態

0)辞書的使用と 6)知識の体系化への利用を除き、3 章であげたすべての利用形態についての研究事例が多く見られ、オントロジー利用法の多様性が伺える。

オントロジーの利用法と利用形態の相関

表の左上から右下に向けた対角線上に沿って、多い件数の組合せが見られることから、より高度な処理を行おうとするにつれて、複雑な定義を含むオントロジーが必要とされていることが分かる。

なお、機能概念オントロジーに基づくセマンティックアノテーションを用いた技術文書管理システム[来村 05, 鷲尾 06]のように 3)意味制約を含むオントロジーを 0)共通語彙として利用した例もある。このことが示唆しているように、複雑な定義を含むオントロジーを基本的な知識処理に用いることにより、システムの高度化を実現することも可能である。

応用分野の広がり

表の数字の中には現れてはいないが、それぞれ研究事例が対象としている分野は、ソフトウェア、Web システム、ビジネス、設計、材料、バイオ、ゲノム、医療、教育、法律...など、非常に多岐にわたった分野で利用されていることが見て取れた。

5. まとめ

本論文では、様々なオントロジーの応用研究を、用いているオントロジーの種類と利用形態という指標を用いて分析し、その傾向について考察した。この指標は、決して完備性のあるものではないし、分析対象とした文献数やそれらの出典も十分な量をカバーしているとは言いがたい。しかし、多種多様な領域で増加しているオントロジーを利用した応用研究の概要を把握することにはできたと考えている。

一口にオントロジーの応用研究と言うのではなく、それぞれの研究の立場を明確にすることで、それぞれの特徴を活かすことができればよいと思われる。実際に、対象とする分野によってオントロジーに求められている役割は大きく異なる。いわゆるライトウェイトなアプローチが効率的な場合もあれば、ヘビーウェイトなアプローチを積み重ねることにより価値の高い成果を期待出来る場合もある。本論文で述べた指標によりそれらの違いを明らかにすることで、オントロジー利用研究の更なる発展に貢献することが期待される。

今後、同様な分析を継続して続けていくとすれば、対象範囲を拡大すると共に、各指標を用いるときに分類基準を再考する必要があると思われる。例えば、3 章で述べた利用形態について、どちらに分類すべきか迷う例もしばしば見られ、より正確な分類するには何かしらのガイドラインが必要と感じられた。また同一の利用形態に分類した研究の中にも、様々な手法を用いたものが含まれており、それらを踏まえた分類の詳細化も考えられる。またオントロジーの種類に関しては、文献にはオントロジーの一部が例示されているだけのものが多く、どの種類に分類するか

の判断が難しいものもあった。研究成果を共有しオントロジー応用研究の発展を促進するという点においては、利用したオントロジーを公開する場の重要性が増すと考えられる。

参考文献

- [Guarino 99] N. Guarino, C. Masolo, and G. Vetere, 'OntoSeek: Content-Based Access to the Web', IEEE Intelligent Systems, 14(3), 70-80, 1999
- [林 03] 林雄介, 山崎龍太郎, 池田満, 溝口理一郎: オントロジーウェアな学習コンテンツ設計環境, 情報処理学会論文誌, Vol. 44, No. 1, pp. 195-208, 2003
- [川村 05] 川村 隆浩, 長谷川 哲夫, 大須賀 昭彦, Massimo Paolucci, Katia Sycara: "セマンティック Web サービスマッチメーカーの公開実験に基づく評価", 人工知能学会論文誌, Vol. 20, No. 6, pp.426-436 (2005)
- [来村 02a] 来村徳信, 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, 人工知能学会誌, 17(1), pp61-72, 2002
- [来村 02b] 来村徳信, 笠井俊信, 吉川真理子, 高橋賢, 古崎晃司, 溝口理一郎, 機能オントロジーに基づく機能的知識体系的記述とその機能構造設計支援における利用, 人工知能学会誌, 17(1), pp.73-84, 2002
- [来村 05] 来村 徳信, 鷲尾 尚哉, 他: 技術知識管理のための機能に関するオントロジーとセマンティックアノテーション, 第 19 回人工知能学会全国大会論文集(JSIAI05), 2005
- [Kitamura 06] Kitamura Y.: Roles of ontologies of engineering artifacts for design knowledge modeling, The 5th International Seminar and Workshop Engineering Design in Integrated Product Development (EDIProD 2006), 21-23, September 2006, Poland, pp.59-69, 2006
- [小路 04] 小路 悠介, 来村 徳信, 他: オントロジーに基づく統合機能モデルとその知識変換への利用 -FMEA シートへの変換-, 人工知能学会全国大会論文集(JSIAI04), 2004
- [古崎 06] 古崎晃司, 飯田哲也, 来村徳信, 溝口理一郎: 軽量オントロジーを用いた簡易型 Web コンテンツ管理システムの開発, 第 20 回人工知能学会全国大会論文集(JSIAI06), 3B2-1, 2006
- [黒橋 01] 黒橋 禎夫, 日笠 亘: 京都大学附属図書館における自動レファレンス・サービス・システム, 情報管理, Vol. 44, No. 3, 2001
- [溝口 05] 溝口理一郎: オントロジー工学, オーム社, 2005
- [大久保 06] 大久保 公則, 小路 悠介, 来村 徳信, 溝口 理一郎: 異なる機能語彙体系間の相互運用性に関する検討, 第 20 回人工知能学会全国大会論文集(JSIAI06), 3B2-2, 2006
- [岡田 04] プロセスプラントとオントロジーパラダイム - EPISTLE データモデルの紹介 -, 人工知能学会誌, Vol.19, No.2, pp.187-193, 2004
- [PSL] <http://www.mel.nist.gov/psl/>
- [高井 05] 高井 貴子, 溝口 理一郎: "デバイスオントロジーに基づくシグナル伝達パスウェイの統一的記述枠組みの開発", 人工知能学会論文誌, Vol. 20, No. 6, pp.406-416 (2005)
- [吉岡 00] Knowledge Intensive Engineering Framework (KIEF) 知識集約型工学のためのフレームワーク, <http://mhjcc3-ei.eng.hokudai.ac.jp/~yoshioka/KIEF/>
- [鷲尾 06] 鷲尾 尚哉, 小路 悠介, 他: Funnotation: 技術文書共有のための機能アノテーションとその利用の枠組み, 第 20 回人工知能学会全国大会論文集(JSIAI06), 3G2-1, 2006

¹ http://www.ei.sanken.osaka-u.ac.jp/hozo/onto_apps/