

# オントロジー研究の基礎と応用

## Ontology: foundations and applications

溝口理一郎

Riichiro Mizoguchi

miz@ei.sanken.osaka-u.ac.jp

大阪大学 産業科学研究所

ISIR, Osaka University

Abstract: I discuss ontology and ontological engineering which are expected to provide us with a solid foundation of so-called Content-Directed AI Research. I first clarify the basic issues of content-directed research by answering typical questions and then describe some definitions of an ontology. I next answer further questions about an ontology and describe the utility of ontology followed by the current state of ontology research and development. Finally, I discuss my idea about Ontological Engineering as well as its future.

Keyword: オントロジー, オントロジー工学, 知識ベース

### 1. まえがき

知識処理/AI の分野で大きなパラダイムシフトが起こりつつある。それは次の三つ、(1) 処理中心から情報中心へ、(2) コンピュータ中心から人間中心へ、(3) 形式中心から内容中心へ、に集約される。初めの二つの動向の重要性もさることながら、21世紀に到来する「知識の時代」に最も深く関係する「形式中心から内容中心へ」の傾向は筆者が主張する「内容指向」研究に関するパラダイムシフトである。これまでのAI研究はフォーマリズム指向性が強く、形式的で、一般性が高い研究が中心的な位置を占めて来た。言い換えれば、「入れ物」を作ることに関わる基礎研究は精力的に行なわれて来たが、何を入れるかという「内容」の議論はその個別性故にないがしろにされてきた。現実の問題を扱うには、対象に関する深い理解とそれに基づく様々な工夫が必要となるが、その議論が欠落しているのである。

知識ベースシステムが独立した自動問題解決器であることを止めて、上述のパラダイムシフトに順応し、オントロジー工学で武装して生まれ変わりつつある。本稿ではそのような動向を総括する内容指向研究の基礎理論と基盤技術を与えるオントロジー(工学) [溝口

97]の基礎と応用についてわかりやすく解説する [JIPDEC98]。

### 2. 基本的な問い

#### 2.1 なぜ知識ベースではだめでオントロジーなのか?

知識は領域依存であり、特に専門家の経験的知識を直接考察の対象にした知識工学は行き詰まり感があることは事実である。しかし、オントロジーは知識を構成する基本概念に立ち戻って考察する。知識の元になる対象世界を客観的な存在として考察し、知識の階層性、知識の分解可能性、そしてコンテキスト依存概念の同定と除去などを注意深く考察することによって、物事や対象の成り立ちを基本から検討し、知識ベースが持つ様々な問題を回避することの糸口を与える。

#### 2.2 内容指向研究は応用研究にすぎないのでは?

確かに内容指向研究は領域固有の知識を対象にする。しかし、領域知識無しにはいかなる知的システムも実際に意味のある振る舞いをなしえないことを再確認することが必要である。内容指向研究は「オントロジー」を武器にして、階層化された適切なサイズの領域内で

「知識を積み上げる」理論の確立を目指している。それはAI研究の実世界への貢献を放棄しない限り避けることができない、実世界の知識を扱うための基礎理論から技術までを対象とする研究であり、従来の基礎と応用という単純な2分法的考えをうち砕く試みである。

### 2.3 内容指向研究は形式を軽視するのか？

内容指向研究は「形式指向」偏重傾向を是正する試みであって、決して「形式」技術を排斥しない。むしろその力は物理学における数学のように不可欠である[池田 98]。その顕著な例がオントロジー定義における公理の利用である。内容指向研究はオントロジー工学を核として内容と形式の適度なバランスをとる。

以下ではこのような背景説明の下に、オントロジーに関する入門的解説を行う。

## 3. オントロジーとは

内容指向研究の核理論・技術としての「オントロジー」は重要な概念であるにも関わらず、それが何を意味するものであるかということに関して多くの混乱が見受けられる。実際、オントロジー関係の会議でしばしば議論されるが合意が得られてはいないのが現状である。ここでは、いくつかの定義と筆者なりの総括を紹介する。

### 3.1 いくつかの定義

(1) 本来、オントロジーは哲学の用語であり、「存在に関する体系的な理論(存在論)」という意味を持っており、世の中に存在する全てのものを系統立てて説明することを目指している。なにが存在として本質的で、なにが存在物の区別をするのか？存在物の間の関係は？などと言うことを考え、その結果としての理論と体系がオントロジーである。

(2) 人工知能の立場からは、「概念化の明示的な規約」[Gruber]という定義がなされる。ここで概念化とは、対象(世界)に関して興味を持つ概念とそれらの間の関係とを指す。我々が世界を認識し、そのモデルをコンピュータ内に作るうとするときには、必ず世界を概念化する。例えば、ビジネスプロセスのモデルを構築する際には、ビジネス自体を構成する概念を洗いだし、それらの間の関係を整理し、概念を特徴付け、他と区別する属性を明確にする。そして、それらの概

念を記述する。概念のなかには、オブジェクト(名詞に対応)だけではなく、活動を表す動詞的なものもある。そして、活動が対象とするオブジェクトとそれらとを関係付ける制約がある。これら全てを体系化したものがオントロジーである。

(3) 知識ベースの立場からは、「人工システムを構築する際のビルディングブロックとして用いられる基本概念/語彙の体系(理論)」[Mizoguchi 93]という定義がなされる。知識ベースは対象とする世界における問題解決モデルを構築するときには不可欠となるものである。問題解決を対象とするので、オントロジーは問題解決過程に固有の概念化であるタスクに関するオントロジーであるタスクオントロジー[ティヘリノ 93][Mizoguchi 95]、そしてタスクが実行される領域(ドメイン)に関わるドメインオントロジーの大きく2種類に分かれる。

#### (4) 緩い定義

「ある目的のための世界の認識に関する共通の合意」by 武田(奈良先端大)。ここでは、定義の一般性を増すことに留意して意識的に合意内容を規定することは避けている。もう一つの特徴はオントロジー目的依存性を明示していることである。

#### (5) もう一つの定義 [Gruber]

オントロジーは複数人の中で共有される合意内容である。共有される合意内容としては、ドメイン知識のモデル化の概念フレームワーク、相互に交信し合うエージェントの会話の内容に関わる通信規約、ドメインの知識を記述するときの合意事項、などが含まれる。知識共有の考えのなかでは、オントロジーは表現のための語彙の定義の形で記述される。最も簡単な場合としては、カテゴリーとそれらの間の上下関係(包含関係)を規定したもの(概念階層)であり、関係データベースにおける概念スキーマがある。

要するに、オントロジーとは『情報処理が対象とする世界のモデル構築者がその世界をどのように「眺めたか」、言い換えるとその世界には「何が存在している」と見なしてモデルを構築したかを(共有を指向して)明示的にしたものであり、その結果得られた基本概念や概念間の関係を土台にしてモデルを記述する事ができる』という理解が有用であると思われる。

### 3.2 典型的な問い

ここではFAQの形式を踏襲して、典型的な質問を想定し、それに答えていこう。

### (1) オントロジーと知識ベースは何が異なるのか？

以下の解答は Ontolingua mailing list において Stanford 大学の Adam Farquhar <axf@HPP.Stanford.EDU>によって、Wed, 26 Feb 1997 に掲載された意見の翻訳である。

- (a) それはその分野の人々（エージェント）の合意した知識を表現したものか？
- (b) 人々はそれを正確に定義された用語として参照しているか？
- (c) その表現に用いられている言語は人々が言いたいことを表現するのに十分な表現力を持っているか？
- (d) それは複数の異なった問題解決に再利用可能か？
- (e) それは安定しているか？
- (f) それは、新しい知識ベース、データベーススキーマ、そしてオブジェクト指向プログラムなどの複数のアプリケーションプログラムを開発するための出発点として利用可能か？

Yes の解答が多ければ（強ければ）多いほど、それだけオントロジーらしくなる。

このように、「オントロジーと知識ベースには明確な境界はない」というのも一つの立場である。それは、オントロジーも言うまでもなく、知識の一種であることから当然の帰結であるとも言える。しかし、「知識ベース」の意味がきわめて曖昧であるのでその定義をすることも必要であろう。例えば、Cycのような知識ベースを意味するのであればこの間はまさに的を射た問であるといえる。しかし、専門家の経験則を蓄積したエキスパートシステムのルールベースを想定すれば、オントロジーとの比較は問題外となる。なぜならそのようなルールベースは開発者が持つ（暗黙的な）オントロジーに基づいて構築されているからであり、決してオントロジーを表現していないからである。オントロジーは「ある知識ベースの暗黙の概念化（対象世界に関する理解）を明示化したもの」であり、知識ベースとはメタな関係にあると考える方が健全であろう。

### (2) オブジェクト指向方法論、あるいはクラス階層と何が違うのか？

確かにオブジェクト指向言語のクラス階層とオントロジーの概念階層との類似性は高い。また、オブジェクト指向方法論はその上流における分析ではオントロジー開発方法論との共通点は多い。しかし、実装に近

いフェーズでは前者は実行性能を重視し、後者は宣言的記述に基づき明示性、形式性を重視する。そして、決定的な相違点は公理の「宣言的」かつ「明示的」記述性の有無にある。すなわち、オブジェクト指向言語は本質的に手続き的であるためクラスの意味、クラス間の関係、そしてメソッドの持つ意味はその手続き的記述の中に埋もれており暗黙的である。一方、オントロジーでは各概念の意味定義と概念間の関係記述は宣言的、かつ明示的である。手続きに関してもその仕様は宣言的に記述され、形式性、明示性は維持される。

### (3) 何が新しいのか？Taxonomy と何がちがうのか？

オントロジーは「Taxonomy（概念階層）」という意味を含意するが、それ以外の意味も含んだ概念である。一般に、新しい用語・概念は既存の概念を発展させたものが多く、全く新しい概念であることはまれである。多くはそれまでに存在していた概念を拡張したもの、あるいは新しい観点から再評価したものなどである場合が多い。オントロジーもその例外ではない。すなわち、従来から存在していた「概念階層」、「標準語彙」、「Upper model」などの概念を含むとともに、知識ベースの開発に際してなされている暗黙の仮定、前提となっている対象世界の概念化（利用目的に依存した概念とそこで興味のある概念間の関係）を述語論理などを用いた明示的かつ形式的な定義という新しい概念が付加されている。

### (4) 汎用性の高い上位レベルオントロジーとタスク依存のオントロジーとは両立するか？

工学的にはオントロジーはその利用目的に依存するものであって、汎用のオントロジーは使いものにならないという考えが支配的である。その代表的な考えがタスクオントロジーであり、問題解決に用いられるオントロジーはタスクが要求するように組織化されているべきであるというものである。しかし、一方では Upper モデルと呼ばれる汎用のオントロジーが存在する。両者の距離は遠い。まず思想の相違は一見決定的であるように見える。しかし、両者の用途を考えると、Upper モデルは領域やタスクを越えた多くの人々の間で共有することを目指した概念化であり、タスクオントロジーに現れるタスク固有の概念においても仮定される基本的なものである。実際、両者の遠い距離を埋めた例は少ないが、不可能であるとは思われない。タスクオントロジーは合意可能な Upper モデルを前提としてその上に積み上げる形で設計すべきなのである。

## (5) オントロジーは単なるラベルの集合か、あるいは実行可能な概念定義も意味するのか？

確かに「オントロジー」が意味するところは広範囲にわたっている。その計算論的意味は極めて幅広い。そこで筆者はオントロジーに計算論的意味づけの観点から次の3つのレベルを設定することを提案している。

### ・レベル1

オントロジーの基本的な機能は、対象世界に存在する概念の切り出し（選択）とそれらの関係の記述である。最も一般的で簡単な記述が階層関係の記述であり、そこには概念のラベルと階層記述だけが存在する。このレベルのオントロジーを最もプリミティブなものであるという意味で、レベル1オントロジーと呼ぶ。Internet上のドキュメントのメタデータ記述に使われるXMLタグ（後述）や探索エンジンの概念階層はこのオントロジーの典型である。

### ・レベル2

次に各概念の意味定義（制約）や関係の記述（公理的記述）が加わることによって、オントロジーを利用したモデル構築において種々の適切なガイドや示唆を与えると共に、オントロジーを用いて記述できるものの全体の性質(Competence)に関する質問に答えることができる。これをレベル2オントロジーと呼ぶ。多くのオントロジーはこのレベルを目指している。

### ・レベル3

このレベルのオントロジーはオントロジーを用いて構築されたモデルのある問題解決における実行のパフォーマンス(Performance)、即ち、オントロジーを用いて記述したものが実行されたときの振る舞いに関する質問に回答できる。このようなパフォーマンスに関する質問に対しては、本質的に手続き的な記述が必要な場合があり、形式的な公理と証明系では答えることができないことが多いのが現実である。そこで準公理が必要となる。KADSのオントロジー[Breuker94]やタスクオントロジー[テヘリノ93][瀬田98]はこのレベルのオントロジーである。

## (6) 多くの人の合意が得られるのか？

この問題は安易な解があるとは思っていない。しかし、合意が得られる「何か」が不可欠であるという信念を共有することは必須である。人間同士の意志疎通が可能であるのはオントロジーを共有しているからである。もう一つの重要な観点は、一見合意形成が困難

に見えても、お互いが依って立つ基本概念体系としてのオントロジーを明示化しながら相互理解を深めて、最終的に合意結果をオントロジーとして得ることができるという逆転の発想は重要である。オントロジーはそのような合意形成の媒体として利用できる。

## (7) 具体的にどのようなメリットがあるのか？

オントロジーの目的依存性から、オントロジーはその利用目的に合致した様々な機能を持っている。ここでは、それらの機能を数え上げる。

### (a) 共通語彙の提供

対象とする世界を記述する際には厳密に定義された、関係者の合意に基づく「語彙」が必要であるが、オントロジーはそのような標準的な語彙を提供する。

### (b) 暗黙情報を明示化

ある領域において仕事をしている人々にとって無意識に仮定しているものが多くある。その典型が日頃頻繁に用いている用語（概念）の定義であり、それらの概念間に存在する基本的な関係や制約である。また、知識ベースは何らかの概念化に基づいているが、その概念化に関する情報は多くの場合暗黙的である。オントロジーはまさにこのような暗黙知識を記述したものであり、それらを明示化する役割を持っている。

ここで注意しなければならないことは、「単に」明示化しただけである、とは思わないことである。この「明示化」がいかに重要であるかは経験した人にしか理解できないと言う宿命的なことがあるが、暗黙の前提を明示化できなかった（しなかった）ことが知識共有・再利用を妨げていた大きな原因の一つであることを認識することは重要である。明示化の効果は計り知れないものがある。

### (c) データ構造

データベースの概念スキーマはデータベースのオントロジーであるが、その意味においてオントロジーは対象とする世界に存在する概念や情報を記述するデータ構造を提供する。

### (d) 知識の体系化

知識を体系化するにはまず厳密に定義され、関係者が合意する概念（用語）が必要である。そしてそれらを用いて様々な現象、観測事象、興味ある対象を説明する理論が記述され、組織化される。このようにオントロジーは知識を体系化する際の拠り所となるバックボーンとしての機能を持っている。実際、筆者らは日本学術振興会未来開拓学術推進事業：シンセシスの科

学「人工知能による協調的シンセシスの方法論」において、設計知識の体系化をタスクオントロジーの概念に基づいて研究を進めている[シンセシスの科学]。

**(e) 標準化**

今後の知識処理の高度化，高能率化を考えると、標準化の問題を避けて通ることはできない。工業製品の生産性の高さを考えれば明らかなように、必要な基本部品の標準化は生産性の高度化に貢献するところが大きい。ボルトとナットに対応するような規格品を知識処理の世界でも作る必要がある。オントロジーは標準概念の意味を規定するものであり、オントロジーの一般性，共通性，形式性は標準化に貢献する。

**(f) 設計意図**

オントロジーはシステム設計者が持っている対象に関する理解，システム設計の意図，すなわち design rationale を明らかにする。オントロジーは前提とされている条件や環境，解くべき問題が要求する仮定などの暗黙的な情報，そしてそれを反映した対象の世界の概念化に関わる根本的な情報を明示し，知識ベースの構築を支えるバックボーンとして機能する。具体的には，診断システムであれば診断可能な故障原因のクラス，推論可能な範囲，定性推論であれば導出可能な因果関係のクラス，ドメインのモデルであれば想定したタスク，など設計者の意図が明らかにされる。

**(g) メタモデル的機能**

モデルは現実の対象を抽象化してコンピュータ内部に作られる。そして、オントロジーはある対象をモデル化するときに必要な概念とそれらの間に成立する関係を明示的に規定し，そのモデルはオントロジーが提供する概念と制約の下で作られる。この意味で、オントロジーはメタモデルとしての機能を持っていることができる[Ikeda 97]。

**(h) Theory of content (内容の理論)**

以上のことを総合して、オントロジーは、ともすれば個別の状況に依存したアドホックな議論になりがちで形式理論の様に積み重ねが効くような方法論を持たなかった「内容指向研究」に対して、内容を扱うために必要な新しい「理論」, Theory of Content[溝口 97][Chandra 99][溝口 99b] を提供する(5.1 参照)。

オントロジーに基づいて開発されたシステムが持つと期待される利点としては、(1)システムや各モジュールに基づいている仮定や前提が明らかにされるので再

利用性が増す。(2)システムの機能や通信プロトコルの記述などが容易になるので相互運用性が増す。(3)対象世界の概念化とその形式的記述はデータの論理構造や概念(オブジェクト)の機能の形式的仕様記述としての役割を果たすため実装容易性が増す。(4)ツールはそれが構築するシステムが持つべきモデルに関するモデル(メタモデル)を持っていることが望まれるが、オントロジーのメタモデル機能は何を生成するためのツールであるかということを知っている知的なツールの構築を可能にする[Ikeda97]。

**4. 研究の状況**

本章では内外におけるオントロジー研究の現状を概観する。

**4.1 理論的研究**

情報科学の立場からのオントロジー理論は N.Guarino と J.Sowa 等によって積極的に進められている。Guarino は FOIS98(Formal Ontology in Information Systems) と呼ばれる、KR98 の直後に開催された世界で最初のオントロジーに関する国際会議の議長であり、Sowa は Conceptual Graph [Sowa 99] の提案者でありオントロジー理論と並行して知識表現の中間言語標準化を積極的に進めている。

両者に共通する思想は、工学的な応用を常に意識しつつ、哲学の研究成果としての Top-level オントロジーを重視する点である。多くの実践派の人々は Top-level オントロジーの有用性を疑問視する。所詮オントロジーは特定の目的に適合するように設計されていなければ有用性の点で問題が起こり得るからである。しかし、大規模な知識ベースをある特定のドメインで構築することを最終目標としてオントロジーを開発する場合には、その妥当性を特定タスクを越えたある程度の一般性があるところで求めなければならなくなり、その際のよりどころとなるのが Top-level オントロジー

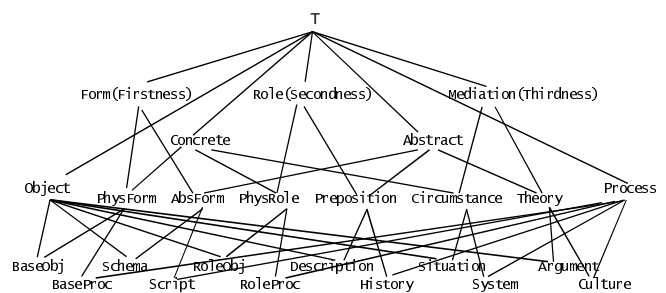


図1. Sowa のオントロジー [Sowa95]

であろう。哲学分野の成果は考察に十分値すると筆者は考えている。

Soawa の Top-level オントロジーを図 1 に示す。その本質は Peirce が提唱したものを少し脚色した、Firstness(それ自身で定義可能な概念:人間、鉄など)、Secondness(他の存在物を参照しなければ定義できない概念:妻、母など)、Thirdness(Firstness と Secondness を関係づける仲介役を演じる環境や状況としての概念。例えば、婚姻、母性など)に加えて、abstract と physical、そして Continuant(実体を持つべき性質)と Occurrent(プロセスを持つべき性質)の二つを追加して、合計7つの基本Propertyを元に12個のTop-level カテゴリーを求めている。

Guarino のオントロジーは更に本格的である [Guarino 97]。全体・部分の理論(部分とは何か、どのような性質を満たせば部分であることを認定できるのか、何種類の部分があるのか、全体とは何か、どのようなときに全体性は失われるのか、などに答える理論)、identityの理論(実体は identity を維持したままどのような変化が許されるのか、どのような時 identity は失われるのかなどの問いに答える理論)、依存関係の理論(依存するとは何か、どのような種類の依存関係があるのかなどに答える理論:上述の Firstness, secondness などの考えに深く関係する)の3つの哲学的理論の成果を踏まえつつ、厳密な Top-level オントロジーを設計している。

もう一つの顕著な特徴は、オントロジーを我々人間が認知する実体を表す Particulars(物体、出来事、物質など)とそれらについて語るときに必要な概念 Universals(概念、属性、状態、関係など)に2分している点である。そして各々についてオントロジーを設計している。ここで中心となっている概念認識の原理は、identity が matter(物質)のように外延的(部分の単なる集合というだけで identity が認定できる)か person の様に内包的であるか、永続性か一過性のものか(Rigidity)などの根本的属性の有無である。

- particulars
  - substrate
    - location
      - space
      - time
  - matter
    - iron
    - wood
    - etc.
- object
  - concrete object
  - continuant

- singular
  - configuration
  - singular body
    - physical body
      - topological body
      - topological whole
        - a planet
      - topological piece
        - a piece of wood
    - morphological body
      - morphological whole
        - a cube of wood
      - morphological piece
        - a mountain
    - functional body
      - functional whole
        - artifact
      - functional piece
        - artifact component
        - organic structure
  - singular feature
    - physical feature
      - topological feature
        - the boundary of blob
      - morphological feature
        - the top of mountain
        - a hole on a piece of cheese
      - functional feature
        - the head of a blob
        - the profile of a wing
  - plural
    - plural body
      - physical plurality
        - mereological plurality
        - topological plurality
          - a lump of coal
        - morphological plurality
          - a line of trees
          - a pattern of stars
        - functional plurality
          - a disassembled artifact
      - plural feature
        - a hole of a lump of coal
        - the sword of orion
    - occurrent
      - situation
      - physical occurrent
        - topological
        - biological occurrent
        - intentional o occurrent
          - morphological occurrent
        - functional occurrent
          - action
          - mental event
          - social occurrent
            - communicative event
      - abstract object
      - quality
        - the color of this rose
        - the style of this talk

図2 Guarino の Particulars のオントロジー [Guarino 97]

図2は Particulars のオントロジーの一部である。大きな特徴は、ともすると自然言語に現れる語彙の範囲で抽象度の高い概念を基準にオントロジーを作っていく傾向がある中で、上述の3つの理論に則って明確な根拠の元に概念選択を行っていること、特に identity 理論と依存関係理論の成果の利用が顕著であり説得力に富む。しかし、単体と複合体を区別してい

るところに問題が見られる。

#### 4.2 機械可読辞書(MRD)

自然言語処理には大規模な機械可読辞書が必要となるため、古くから辞書の開発が活発であり、情報処理の分野の中では早くからオントロジーの議論が行われていた。前述の定義から辞書とオントロジーが同一であるとは言えないが、最近の MRD はその上位概念の構造を Upper model と呼んでオントロジーと同列に議論ができる様な構成をとっている例が多い。その代表的なものが、WordNet [Miller93][WordNet], EDR [横井 97], EuroWordNet [EWN], Generalized Upper Model [GUM], などであろう。辞書ではないが自然言語処理を意識して構築された常識ベース Cyc [Cyc] も忘れてはならない。ここでは詳しい議論はできないので、詳細は JIPDEC の「オントロジー工学委員会」調査報告書 [JIPDEC98] を参照されたい。

ここで言及しておきたいことは、ANSI ad hoc 委員会にオントロジーの標準化を行う委員会ができており、すでに 2 年ほど活動が続けられていることである (URL: <http://ksl-web.stanford.edu/onto-std/>)。実現するかどうかは別として、Cyc, WordNet, EDR などの知識ベースや機械可読辞書の Upper model を取り出してそれらの共通性、相違点を洗い出して概念間の対応をとりながら Reference Ontology (RO) と呼べるようなものを作るべく努力を重ねている [荻野 98]。今のところ、相違点の方が共通性を圧倒しており RO が無事構築される可能性は大きくはないが、とにかく彼らは果敢にもオントロジーの標準化という困難な問題に取り組んでいることは注目に値する。

#### 4.3 Metadata, XML タグ, そして知的 Player

標準化に関わるもう一つの動きは Dublin Core [DC97], MCF: Meta Content Framework [MCF97] や W3 コンソーシアムの RDF: Resource Description Framework [RDF97] に代表される Internet 上の情報を対象にした Metadata 記述に関する標準化活動である [浦本 98]。Dublin Core はそもそもメタデータとしてどのような情報を記述するべきかという基本的なことに関するワークショップが開催されそのときの成果をまとめた物であるが、de fact 標準になりつつある。MCF は Netscape 社が W3 コンソーシアムに提案したメタデータ記述のデータモデル、構文、語彙などの標準化案であ

る。構文は XML を採用している。RDF は MCF 等を元に W3 コンソーシアムで標準化されつつあるメタデータ記述のフレームワークである。

XML タグセットは特定領域の人々の間で共有することを前提として、対象ドキュメントの意味に踏み込んだ記述を行うのが普通である。その応用として、例えばカルテや看護記録の標準化と共有等が行われている [看護サマリーの電子化]。また、知的教育システムでは学習者の理解状態に応じて単元の進み方、問題の提示、説明の提示などを適応的に変化させる必要があるが、そのようなシステムの振る舞いを制御することに関わる情報と制御対象を識別するタグなどを開発・共有することによって、XML タグ付け教材を作成することができる。そして、レベル 2 のタスクオントロジーを適用することによって、そのような XML 教材を解釈実行するエンジンの仕様をも標準化することができる。そのようなエンジンは「Plug & play の概念に基づく知的な Player」と呼ぶことができるが、今後は教育に限らず多くの分野で XML タグとタスクオントロジー、そして Java で実現された知的 Player との関わりは深まるものと思われる [AIED99 WS]。

#### 4.4 エージェントとオントロジー

マルチエージェント環境においてエージェント間の通信を可能にするためには共通のオントロジーを持つことが必須である。FIPA: Foundation for Intelligent Physical Agents はオントロジーにアクセスするためのエージェントの標準化を試みている [FIPA 98]。

#### 4.5 オントロジー開発の具体例

オントロジー開発の例は豊富になっている。まず、Stanford 大学の Ontolingua を使って開発されたオントロジーベースがあげられる [Ontolingua]。そこには 50 ほどの開発例が蓄積されている。ビジネスプロセスモデリングやエンタプライズモデリングの分野でのオントロジー開発も盛んである [伊藤 98]。特に、エディンバラ大学の AIAI 研究所 [AIAI] とトロント大学 [TOVE] で開発されたエンタプライズオントロジーは特筆に値する。共に、開発の過程に関する論文もあり [Uschold 98] [Gruninger 95]、参考になる。最新の活動としては IEEE Intelligent Systems (旧 IEEE Expert) のオントロジー特集記事も情報量が多い [Swartout 99]。

オントロジーの情報検索における概念階層の利用は

誰しも思いつくことであるが、実際その利用例は多数ある。その中で一つだけ SHOE プロジェクトを簡単に紹介する[Heflin 98]。SHOE は XML などと同様 Markup 言語であるが、オントロジーを直接記述したり、オントロジーの解釈をするルールまでも書く (Markup する) 機能を持つことが特徴である。その流れは OML (Ontology Markup Language) にまで発展している[OML]。ISO の STEP: Standard for the Exchange of Product model data の新しい活動 ISO-15926 では EPISTLE Framework で提案された包括的なオントロジーが元になっている[EPISTLE]。

我が国での開発例もある。静岡大学の山口等による法律オントロジーの開発は後述の独自の構築支援ツールを法律の専門家が利用して構築された物である点が特長である[山口 98]。奈良先端大の武田等は知識コミュニティとオントロジーとの関わりの研究を行っている[武田 99]。東京大学の富山等も長年の知的 CAD の研究の成果である KIEF のオントロジーを抽出して見通しの良いシステムの記述・構築を行っている[関谷 99]。

学習支援システムにおけるオントロジー開発も盛んに行われている[Mizoguchi 96][Ikeda 97][AIED99 WS]。特に、オーサリングシステムはオントロジーのメタモデル機能を活かす格好の応用対象である。実際、筆者らは訓練タスクオントロジーに基づく、訓練システム教材開発のためのシステム SmartTrainer/AT[Ikeda 97][Jin 99]を開発している。この話題では教材の再利用という大きな課題がある。特に最近の Web-based の学習支援が更に普及することを予見してその再利用、円滑な流通がホットな話題となっている。そのために、IEEE において学習教材メタデータを中心とした標準化活動が精力的に行われている[IEEE LTSC]。

通産省産業科学技術研究開発制度「ヒューマンメディア」プロジェクトの一つ、「次世代プラント用ヒューマンインタフェース」において石油精製プラント運転支援のための本格的な「石油プラントオントロジーとオントロジーサーバ」、そしてその利用方式が開発されている[ヒューマンメディア 99][佐野 99]。常圧蒸留塔周りの概念約 500 個が組織化され、定義されている。オントロジーサーバは、対象とするプラントのモデル構築と構造情報の提供、運転員へのメッセージ生成における語彙の標準化、エージェント間の通信の共通語彙の提供などの機能を果たしている。

未来開拓学術研究推進事業「人工知能による協調的

シンセシスの方法論」プロジェクトにおいては、シンセシスにかかわる世界を対象として機能概念を中心としたオントロジー設計が行われている。機能概念にまつわる 3 種類のカテゴリ (ベース機能、機能タイプ、メタ機能) の同定、各カテゴリ内での機能概念の階層的組織化、ベース機能概念の振る舞い・構造概念への Grounding、そして対象の機能理解システムの開発などを行い、包括的な機能オントロジーの開発と利用、そしてそれに基づくシンセシス知識の体系化研究が行われている。[溝口 99a]

既知と知っている概念の厳密な定義を試みると、意外と我々はよく分かっていないことに気づいて愕然とすることがある。筆者らが最近行った「故障のオントロジー」の研究はその典型例であろう[来村 99a, 99b]。そこでは、「故障」というありふれた概念の概念化を精密にした後に故障診断システムを再考察した結果、多くのシステムの能力の相互比較ができるだけでなく、新しい能力を持つ診断システムの設計が可能になった

#### 4.6 方法論

オントロジーの開発は容易ではなく、優れた方法論と開発環境が望まれる。ここではいくつかの方法論を紹介する。

##### 4.6.1 IDEF5[IDEF5]

企業活動やシステムのモデリングと分析技術とツールの標準化を主に行っている KBSI (Knowledge Based Systems, Inc.) は企業活動モデリングを意識したオントロジー開発方法論とオントロジー表現言語、IDEF5を開発している。オントロジー開発の基本は、語彙の開発にあることを明確にし、用語の抽出、その厳密な意味定義、そして用語間の意味制約を形式的に記述することを方法論の骨子としている。かなりしっかりしたドキュメントもあり、そこには開発の際の注意事項などもかかっているが、オントロジーの図的表現言語を整備しているところが一つの特徴であろう。詳細は[伊藤 98]に譲る。

##### 4.6.2 TOVE[TOVE]

カナダのトロント大学で開発された TOVE 方法論は IDEF5 と同じく、企業活動モデリングのためのオントロジー開発方法論であるが、ユニークな位置を占めている。基本思想を以下に示す。

- (1) オントロジーに期待する能力を、オントロジーを用いて構築したモデルへの質問(Competency questions)という形で整理する。すなわち、オントロジーに基づいて構築したモデルに答えてほしい質問を要求仕様の形で整理する。TOVE ではオントロジーはモデルを規定する公理系として機能することが前提となっている。
- (2) オントロジーの公理はProlog を用いて記述され、上述の質問に答えることによってオントロジーを検証する。

このように要求仕様を明確にしてオントロジーを設計し、それを評価に用いるというソフトウェア開発の正統的な方法論である。企業活動オントロジー開発のために彼らが用意した質問の例は：

- (1) あるゴールを達成するためにはどのような活動が遂行されなければならないか？
- (2) 未来の複数時点における活動が与えられたとき、それ以外の時点における資源と活動に関する性質を聞く。
- (3) あるタスクを前に（後ろに）移動するとどのようなことが起こるか？

などである。これからもわかるように TOVE では時間と活動の厳密な定義を公理化して、概念レベルにおけるモデルの実行を通して定性的な範囲でのモデルの振舞いに関する様々な性質を調べることができる。

#### 4.6.3 DODDLE, アスペクト理論, CLEPE & AFM

我が国におけるオントロジー開発環境を紹介する。DODDLE( A Domain Ontology rapiD DeveLopment Environment) [Yamaguchi97][山口 99] は静岡大学の山口らによって開発された環境で法律オントロジーの開発に使われた実績がある。一般的な言語知識である WordNet の概念体系を利用して、法律という領域固有オントロジーを開発する際に適切にガイドする機能があることが特徴である。開発されるオントロジーは現状ではレベル1である。

奈良先端大の武田らのアスペクト理論 [Takeda95][武田 99]は同一の世界を複数の視点から見た場合にできる複数のオントロジーを統合する枠組みであり、かつオントロジー構成の方法論を提供している。一つのオントロジーをアスペクトと呼び、それらの間の論理演算を定義して、オントロジーの統合操作を支援する。ASPECTROL という言語が用意されており、

エージェントコミュニケーションにおけるオントロジー記述にも利用されている。対象となるオントロジーはレベル1である。

CLEPE は大阪大学の池田・瀬田らによって開発されているレベル3のタスクオントロジー記述言語・利用環境である[瀬田 98]。3種類のユーザ、すなわち、基本オントロジー開発者、タスクオントロジー開発者、そしてタスクオントロジーを用いてモデルを開発する利用者を想定して言語設計が行われている。モデルはタスクオントロジーの公理の制約の下に作られ、モデルの整合性が維持される。また、モデルは概念レベルでの実行が可能であり TOVE と同様様々な質問に答えることができる。

AFM(Activity-First Method)[高岡 95][Mizoguchi95][久保 99]はタスク依存のドメインオントロジー開発の方法論である。技術ドキュメントからの開発を想定しており、はじめにタスク分析を行って「動詞概念」を抽出し、動詞が目的語にする「名詞」概念をそのタスクにおける「役割」の下に組織化するという手順を踏む。システムはソースドキュメントから数種類の間生成物を経て最終出力のタスクオントロジーまでの変化を管理しており、きめ細かい支援が可能となっている。レベル2のオントロジー開発を対象にしている[久保 99]。またオントロジーをグラフィカルに記述するエディタ「法造」が開発され、研究室内で盛んに用いられている[古崎 99]。

オントロジー構築方法論全般について言えば、ソフトウェア開発支援ツール的一种にとどまっているのがほとんどで、正しいクラス認定や関係の同定などに踏み込んだ支援はオントロジー基礎論の発展を待たねばならないのが現状である[溝口 99b]。

## 5. オントロジー工学

本節ではオントロジー工学そのものに関する具体的な問いに答えてみたい。

### 5.1 「オントロジーは内容の理論である」とはどういう意味か？

物理学が理論科学として見事に成立していることの原因を考えると数学の働きが大きいことがわかる。数学には、

- (1) 自然数、実数、連続関数、解析関数  
などの厳密に定義可能な変量のクラスと、それを操

作したり、関係を記述する

(2) 四則演算, 微分, 群, 環, 線形性, 直交性  
などの演算クラスや関係が定義されている。

そして, 物理学が対象とする現象を説明するために必要な変量と演算のほとんど全てが数学で既に準備されている概念に置き換えることができる。このことが物理学を理論として成立させている根本にある。すなわち興味ある現象という「内容」を「形式的に」扱う理論体系が存在するということである。

それでは知識処理の場合はどうであろうか? 理論の典型である述語論理を例に取る。それを知識表現と見なした場合, 本来豊富であるべき知識の「内容」を, 真理値をとる「述語」という「一つの」カテゴリだけに押し込めて, その見返りとして美しい形式理論を手に入れているといえる。P(X)という述語はそれがあるオブジェクトXに関して真か偽の値をとる「何か」である以外の何者をも意味していない。Xが赤いのか, 重いのか, 物なのか, 人なのか, 学生なのか, 夫なのか, 動いたのか, 歩いたのか... 要するに世界に関して何を主張したいのかは一切捨象されているのである。このことから述語論理は, 物理学における数学の強力な貢献と比べると無力であることが分かる。

オントロジーは述語論理表現に必要な「語彙」, すなわち具体的な述語とそれに関する意味論(意味制約など)を提供する(言うまでもないが, 述語論理はオントロジー無しには世界に関して何も主張することはできない)。世界知識を表すときに必要となる述語を階層的に分類して, 各カテゴリに属する概念(述語)が, 真理値をとる以外にどのような意味や制約を持つのかを規定する。たとえば, tall(X), human(X), student(X), move(X)などの4つの述語は本質的な相違点を持っていること, たとえば, tallは属性でありidentityを持たない, humanはXの本質的属性を表す概念でありidentityを持つこと, studentはrole概念であってXの本質属性を表さず, プロセスを表すmoveと同様一過性のものである, などである。そして, これらの異なった述語はそれらが属するカテゴリに固有の性質を持ち, それに依存する様々な制約や関係を持っている。オントロジーはそれらを明示化し, 形式的に厳密に記述する。すなわち「内容の理論」を提供する。

物理学の進歩は, 上述の様な数学上の概念と演算・関係を使用しながら, 運動学, 熱力学, 電磁気学, などのように個別の分野で知識を体系化しつつ, 各種の

理論の統合も研究してきた成果を「蓄積」してきたことにある。一方, 知識処理の研究者は「内容はあまりにも領域個別性が高い」ということを理由に「蓄積する」ことを放棄してきた。オントロジーに現れる概念階層は, どの抽象レベルまでは「共通の内容の理論」として受け入れることができるかと言う議論も含めて, 内容の理論を構築し, 蓄積することに貢献する。我々は数学の様に数少ない概念と操作だけで閉じることはないことを認識しつつ, 「知識(内容)の理論」構築へ向けて長期的視野を持って取り組まなければならない。

## 5.2 オントロジー工学がカバーする範囲は?

オントロジー工学の研究課題としては少なくとも下記の課題を挙げることができる。

### 基幹研究

- ・哲学(存在論)
- ・科学哲学
- ・知識表現

### 常識ベース研究

#### オントロジー構築方法論

- ・記述言語
- ・方法論, 構築支援環境
- ・オントロジー比較, オントロジー統合
- ・合意形成

#### オントロジー開発

- ・上位オントロジー
- ・各種ドメインオントロジー
- ・タスクオントロジー

### 標準化問題

### 評価方式

#### オントロジー応用

- ・知識共有/再利用方式
- ・Knowledge Management, Corporate Knowledge
- ・EDI/CALS
- ・Business Process Modelling
- ・各種知識の体系化
- ・Internet 情報処理, Metadata

### メディア

- ・メディアオントロジー
- ・知識メディア共通オントロジー
- ・メディア統合

### 5.3 オントロジー工学に方法論はあるのか？

オントロジー構築の方法論は上述のように提案が行われつつあるが、オントロジー工学となるとおぼつかない状況である。オントロジー工学の方法論はこれから衆知を集めて走りながら考えて行くべき物であると考えている。あえて、筆者のこれまでの経験を下にした非常に個人的な見解を述べるならば、

- (1) 当然と思っていたことを詳細に検討すること
- (2) 暗黙の仮定を洗い出すこと
- (3) 概念、特に基本概念の定義の明確化、定式化を試みること
- (4) 概念間の関係に注目すると同時にそれがどのような関係であるかを明確に概念化し、定式化すること
- (5) システムの処理自身も概念化の対象にすること
- (6) 明示化した仮定や概念、関係を用いてシステムを記述してみること

などが挙げられる。

### 6. むすび

本稿では、オントロジーに関する包括的な議論と、海外の動向を紹介すると共に、オントロジーとオントロジー工学について解説した。最後に重要と思われる課題をもう一度論じてみたい。

#### (1) オントロジーの標準化

すでに ANSI における活動は紹介したが我が国における標準化の意識は高くないことに危惧を感じる。知識の共有や interoperability, ドキュメントの検索や再利用などは基本語彙や概念、そして最小限のタグなどを共通化する以外に効果的な方法はないことはほとんど自明ではなからうか。そもそも人間がコミュニケーション可能であるのは語彙を共有しているからである。我が国でオントロジー開発とその分野での標準化に関する実質的な活動が生まれることを期待したい。

#### (2) オントロジーに関する共通の理解

既に述べたように、未だにオントロジーの定義については研究者の間で合意は得られていない。というより、むしろ"オントロジーとは何かという議論はやめよう"という動きがでてきている状況にある。少なくとも、自然言語処理に携わるもの、知識ベースの立場をとるもの、哲学的な立場をとるものの3者では合意を得ることは困難であるように見受けられる。自然言語研究者は"語彙"あるいは"辞書"という観点を重視し、知識

ベース研究者は"実行可能性"や知識ベース構築の"部品"という考えを重視し、哲学者は"共通性、汎用性の高い上位モデル"の立場を固執する。いずれも十分強固な正当性の根拠を持っており議論を収束する方向に持っていくことは容易ではない様に見受けられる。

#### (3) オントロジー工学への道

このような状況ではあるが、オントロジーが重要であるという点では研究者の合意は強固である。オントロジーの厳密な定義には触れないようにして、各自が必要と考えるオントロジー理論、システム、そしてオントロジー自身を開発する事に専念することで今は十分であるようにも思われる。現在はオントロジー研究は幼児期にあり、はじめの一步で躓かないことの方が重要であろう。

そこで重要なことはオントロジーの定義ではなく、オントロジー工学の旗印の下で何をなすべきかということを確認にすることであると考えられる。そして、過度な期待は慎みつつ、挑戦的であることが大切であろう。

謝辞：いつも激論を戦わしている、大阪大学産業科学研究所、池田助教と来村助手に感謝します。JIPDEC オントロジー工学委員会での調査活動と討論は大変有意義であった。ここに委員の方々に謝意を表します。

#### <文献>

- [AIAI] <http://www.aiai.ed.ac.uk/~enterprise/enterprise/ontology.html>
- [AIED99 WS] AIED99 Workshop on Ontologies for Intelligent Educational Systems, <http://www.ei.sanken.osaka-u.ac.jp/aied99/aied99-onto.html>.
- [Breuker 94] J.Breuker and W.V.de Velde: The Common KADS Library for Expertise Modelling, IOS Press, Amsterdam, 1994.
- [Chandra 99] Chasndrasekaran, B., J.R.Josephson and R. Menjamins, Why are ontologies, and why do we need them?, IEEE Intelligent Systems, pp.20-26, 1999.
- [Cyc] <http://www.cyc.com/public.html>
- [EWN] <http://www.let.uva.nl/~ewn/>
- [DC97] Dublin core metadata; [http://purl.org/metadata/dublin\\_core](http://purl.org/metadata/dublin_core), 1997
- [EPISTLE] EPISTLE Framework V2.0, <http://www.stepcom.ncl.ac.uk/epistle/epistle.htm>
- [FIPA 98] FIPA 98 Specification Part 12, Ontology service. <http://www.fipa.org>
- [Gruber] Gruber, T., *What-is-an-ontology?* <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
- [Gruninger 95] Gruninger, M., and Fox, M.S., Methodology for the Design and Evaluation of Ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal, 1995.
- [Guarino 97] N.Guarino: Some organizing principles for a unified top-level ontology, Working Notes of AAAI Spring



e  
v OML 調査 (