

概念化アスペクトがオントロジー構築支援に果たす役割について

About Roles for Conceptualizational Aspect in Ontology Constructing Support

上田 俊夫*¹
Toshio UEDA

山口 高平*²
Takahira YAMAGUCHI

池田 満*¹
Mitsuru IKEDA

*¹ 北陸先端科学技術大学院大学知識科学研究科
School of KnOWLedge Science, Japan Advanced Institute of Science and Technology

*² 慶應義塾大学理工学研究科
Graduate School of Science and Technology, Keio University

In this paper, we propose an ontology description language called CAML for supporting ontology authors with scarce ontology literacy and discuss the roles that conceptualizational aspect (CA) plays in ontology construction. CA provides a basic representation scheme for accumulating and sharing knowledge for discerning role concepts obtained from experiences in ontology construction.

1. はじめに

オントロジー構築の作業は、概ね、「対象世界にある概念を見だし、その概念の組成を峻別し、オントロジー記述言語によって表現する」という3つのサブプロセスからなっている。このうち、特に最初の2つのサブプロセスにオントロジー構築の難しさがあると言って良い。概念の認知と峻別の方法論の整備が進んでいる[溝口 05,06]ものの、その習得は決して容易ではない。概念の認知と峻別に必要な知識・スキルが抽象的で個性が高く複雑であることが、一つの大きな原因であると思われる。そのことが、オントロジー構築スキルの習得は、具体事例の構築経験を共有しながらの徒弟的学習にたよらざるをえないという実情をもたらしている。

オントロジー記述言語による形式化によって、オントロジーの一貫性の論理的検証が可能(OWL の場合は記述論理を基礎とした機械推論により検証できる)になっていることはオントロジーの質を保証するうえで大きな支えになっている。しかし、そのような形式性の高いオントロジー記述言語の処理系・環境が提供する機能は、残念ながら上述の3つのサブプロセスの最後の段階の表現の部分の支援に限定的になっている。記述したオントロジーの問題(バグ)の同定に関して示唆的な情報を与えることはあっても、概念を適切に認知・峻別するための手がかりを与えることはない。

本研究は、オントロジー構築スキルを十分に習熟していない人(オントロジーリテラシーが十分に高くない人)に対して、OWLベースのオントロジー構築環境が提供するような形式的な側面での支援に加えて、内容的な側面での支援を提供する環境を構成することを目的としている。そのためには、オントロジー基礎論の抽象的・包括的・(哲学的)な知識と、事例への依存性が極めて高い構築スキルについて、部分的であっても何らかの明示的な表現を与え、抽象的な知識と個別の事例と結びつける仕組みを作り、暗黙性の高いオントロジー構築の経験知を可能な限り明示して蓄積することが重要であろうと思われる。

このような目的意識のもとで、我々は、オントロジー構築方法論の中で理解が難しく、かつ、オントロジーを適切に構築するうえで重要なルール概念に着目し、ルール概念を認知・峻別する

ための知識に表現を与え、ルール概念の認知・峻別の経験知の蓄積・共用のためのツール群を開発することを目指している。ルール概念を適切に定義するためには、その概念自体(例えば教師というルール)の定義だけではなく、ルールを生み出す概念(例えば、学校という全体概念)を見だし、全体に対するルールの役割等の関係性を明らかにする必要がある。溝口らの研究によれば、この全体に対するルールの依存性・関係性は、極めて多様で複雑であることがわかっている[溝口 05]。彼らの研究グループは、この多様性・複雑性を、オントロジー構築の実践経験を踏まえてオントロジー基礎論の立場から原理的に分析したうえで、そこで得られた知見をオントロジーエディタ法造[Hozo]にパターンとして組み入れてユーザに提供するアプローチを進めている。

これに対して、本研究では、ビルトインの基礎論的知見は最小限(最上流)としながら、

- オントロジー構築コミュニティがオントロジー構築の経験を明示化して積み上げ
- その経験知をコミュニティ自らが汎化・体系化して表現し共有する

プロセスを支えるオントロジー構築環境の開発を目指している。本稿では、その基礎となる概念化アスペクト記述言語 CAML について報告する。CAML の特長は、ルール概念の認知・峻別の意図が定式化され、その定式化に基づいたターゲットオントロジーが CAML 言語処理系によって OWL 記述として生成され、定式化の一貫性が記述論理(DL)推論により半自動的に検証される点にある。

以下、本稿では、概念化アスペクト記述言語 CAML の設計意図とその言語仕様について説明する。まず2節で概念化アスペクト(CA)の概要とオントロジー構築において CA が果たす役割について、次いで3節では CAML にビルトインされる最上流概念であり、経験知を汎化・体系化して共有する基盤である CA メタ定義について概要とその役割について説明する。4節では CAML の処理系を中心としたオントロジー構築環境について説明して、最後にまとめとする。

2. 概念化アスペクト(CA)の役割

本研究では、ルールの全体に対する依存性を(全体概念の観点からルールを見るという意味で)アスペクト依存性と呼び、概

念のAspect依存性を認知・峻別する知識を概念化Aspect (CA)と呼んでいる[上田 07]. CAMLは概念化Aspectを表現し、オントロジー中で定義される概念が、どのような観点から認知・峻別されるかを表現する言語である。ここでは CAML の概要を示しながら、概念化Aspectがオントロジー構築において果たす役割について考察する。

2.1 概念化Aspect(CA)の概要

CAMLの詳細を示す準備として言語設計の意図を簡単に説明する。図1に「人が病院において医師というロールを担う」というロール概念を模式的に示している。このとき、医師は病院のAspectに依存していることを明確にすることが、ロール概念の認知・峻別である。病院のAspectに依存するとき、人は、病院長(医師でもある)、医師、事務職員といったロールを担う。このことが図2では、医療組織-人間 CAとして記述されている。以下で図2を簡単に説明する。

Aspect(概念)

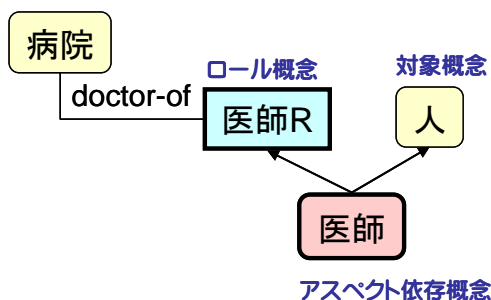


図1ロール概念(医師 R)認知・峻別の模式図

- 1) CA-IDとCA定義型: 全てのCA定義は一意に識別可能なURIをid属性の値としてもつ。また、全てのCAはtype属性に定義型をもつ。医療組織-人間CAの定義型は、組織-人間CAMである。定義型の意味は後述する。
- 2) 対象概念: CAで定義するロールを担う最も典型的な概念(人間)をtarget-concept要素の値として一つ指定する。
- 3) Aspect概念: ロールを生み出す(ロールが依存する)Aspect概念(病院)のうち最も典型的な概念(医療組織)をaspect-concept要素の値として一つ指定する。
- 4) ロール定義: 対象概念が担うロールをrole-def要素と与える。基本的な属性としては、ロールID(id), ロール定義型(type), ロールの成立条件がある。図では、医師はidがca:医師R, ロール定義型がcam:役付きメンバーRM, ロールの成立条件が医師の資格を持つこと(doctor-of)であることが記述されている。ロール定義型については、CA定義型とあわせてあとで説明する。

```

<!-- CA-ID, CA 定義型 -->
<ca id="ca:医療組織-人間 CA" type="cam:組織メンバーCAM">
<!-- 対象概念 -->
<target-concept ref="人間"/>
<!-- Aspect概念 -->
<aspect-concept ref="医療組織"/>
<!-- ロール定義 -->
<role-def id="ca:構成員 R" type="cam:メンバーRM"
rel="member-of"/>
<role-def id="ca:医師 R" type="cam:役付きメンバーRM"
extend="ca:構成員 R" rel="doctor-of"/>
<role-def id="ca:病院長 R" type="cam:リーダーRM"
extend="ca:医師 R" rel="leader-of" cardinality="1"/>
.....
</ca>
    
```

図2 CA 定義例(医療組織-人間 CA)

2.2 基本概念とAspect依存概念

病院の例では、人は病院Aspectに依存しない概念、医師は病院Aspectに依存する概念である。Aspectに依存しない概念を基本概念、Aspectに依存する概念をAspect依存概念と呼ぶことにする。オントロジー工学の基礎論的立場では、何を基本概念として認めるかについて慎重な議論がなされなければならないが、本研究では CAML の言語設計の観点で、次のように問題を簡単化することにする。

ある対象に対してオントロジーを構築しようとするコミュニティにおいて、ある概念の存在が他の概念の存在と独立に認知されると合意できるとき、その概念を基本概念とする。

ここでは、人は基本概念と認知され、その人は家庭というAspectに依存して父・母などのロールを、病院というAspectに依存して医師・事務職員などのロールを担うと認知されると仮定して説明を進めることにする。

2.3 基本概念の階層(基本オントロジー)とCAの展開(Yarn)

CAMLにおいては記述論理による整合性検証機能を活用するために、基本概念の階層はOWLで記述することとしている。基本概念(人間)に対してCA(ca:医療組織-人間CA)を適用してAspect依存概念群(医師, 病院長, 事務職員など)を定義することを、CAを展開すると言う。病院の例では、人に対してca:医療組織-人間CAを展開したことになる。

どの概念に、どのCAを適用するかはCAMLのYarnを用いて記述する。Yarnは、展開するCAを表すref-ca属性、CAの対象概念と基本概念の対応付け記述のassign-target-concept要素、Aspect概念と基本概念の対応付け記述のassign-aspect-concept要素、対象概念の下位概念にロールを対応付けて概念化したAspect依存概念をrole-holder要素に記述する。

YarnによるCAの展開は、デフォルトではCA定義のロールがそのまま展開されるが、Aspect依存概念を詳細化限定することに伴う詳細化展開をすることもできる。図3に、その例を示している。ここでは、展開にあたってのAspect依存概念(assign-aspect-concept)に、図2のCAに記述された医療組織を詳細化限定した“会社医務室”を指定したので、その詳細化限

```

<yarn ref-ca="ca:医療組織-人間 CA">
<!-- 対象概念の対応付け -->
<assign-target-concept class="人間"/>
<!-- 参照概念の対応付け -->
<assign-aspect-concept class="会社医務室">
<!-- Aspect依存概念の指定 -->
<role-holder role="ca:医師 R" class="産業医"/>
<role-holder role="ca:構成員 R" class="医務室職員"/>
.....
</yarn>
    
```

図3 Yarnの記述例

定にもなっており、医師ロール(ca:医師R)はAspect依存概念の産業医に、構成員ロール(ca:構成員R)はAspect依存概念の医務室職員に詳細化展開することを表している。

2.4 概念化Aspect(CA)の役割

既に述べたように CAML の言語設計上の目的は、ロール概念の認知・峻別の経験知の蓄積・共用の基礎となるロール概念認知・峻別の知識に表現を与えることにある。これまでに述べたように、CAMLでは次の3要素:

- OWLによる基本オントロジー記述
- 概念化Aspect(CA)の定義

- 基本概念と概念化アスペクト定義の結合(ルール概念の展開の指定)

を分けて記述するようになっている。この構成は、CAML 設計に際して我々が設定した要求に基づいている。

要求1) オントロジーに含まれるアスペクト依存概念群について、それが、どのようなアスペクトに依存して認知され、個々のアスペクト依存概念が、どのような制約によって峻別されるか記述できるようにすること。

要求2) 概念化アスペクト記述の共用性(一般性)を高めるために、(適切な程度の)抽象化した記述を許し、かつ、その適用(展開)に際しては詳細化した概念を指定できるようにすること(産業医の例)

このように、概念化アスペクトは、アスペクト依存概念の認知・峻別の経験知と具体的なオントロジー構築事例を結びつけて記録するための表現としての役割を担っている。また、さらに重要なことは、その経験知を汎化・体系化するための道具として働きうる点である。序論でのべたように、概念の認知と峻別に必要な知識・スキルが抽象的で個別性が高く複雑であることがオントロジー構築技術の転移性の大きな阻害要因になっていると我々は考えている。もちろん、同じ理由で、概念化アスペクトを、多くの対象をカバーできるように網羅的・体系的に予め整備することは、極めて難しいと考えている。本研究では、オントロジー構築者の分野横断的・持続的な専門家コミュニティ活動(断片的な経験知を蓄積し、可能な限り汎化・体系化する活動)を支える技術としてCAML処理系の開発を進めている。

3. 概念化アスペクトのメタ定義

序論で述べたように、わずかではあるが、オントロジーの最上流概念(生物、組織など)のメタ定義と、基本的な概念化アスペクトの雛形など、基礎論的立場の知見がCAMLにビルトインされている。これを本研究では概念化アスペクトのメタ定義(以下、CAメタ定義)と呼んでいる。厳密にはメタ定義にはなっていないが、2章で説明したCA定義をCAメタ定義のインスタンスと位置づけ、さらにそのCA定義のインスタンスがYarnによってオントロジー中のルール概念が展開されるとい、ランクが違う階層があることを明確にするために、CAメタ定義と呼ぶこととした。

例えば、「組織の構成物は、組織というアスペクトにおいてルールを担う」という上流のCAメタ定義のインスタンスとして「医療組織の構成員は、医療組織において医師・リーダー・事務職員の役割を担う」というCA定義があり、そのインスタンスが「会社の医務室の構成員は、医務室において産業医・医務室長・医務職員の役割を担う」と詳細化されて、Yarnによってオントロジーに結びつけられる。メタ定義は、CAML処理系の前提となる合意レベルであり、処理系が提供する諸機能(編集支援機能・デバッグ機能)はメタ定義に基づいて実現されている。もちろん、CAMLで記述される全てのCA定義はCAメタ定義に整合しなければならず、その整合性は処理系によって検証され、違反は検出されてオーサに通知される。

CAメタ定義は[溝口05]のルール概念とpart-of関係の分析を参考にして検討を進めており、分析結果をCAMLにビルトインしようとしている。

まずロールのアスペクト概念への依存の仕方によって4種類のCAメタ定義を用意しており、それらを基本CAMと呼ぶ。cam:組織メンバーCAMの上位CAMであるcam:外部参照CAMのように対象概念に対して外部のアスペクト概念に依存するもの他、アスペクト概念との相互関係に依存するもの、アスペクト概念の属性の値に依存して定義されるもの、他のアスペクト依存概念に対して相対的に定義される概念を用意している。

```

<!-- CAM-ID, 上位 CAM-->
<cam id="cam:組織メンバーCAM" super="cam:外部参照 CAM">
<!-- 対象概念制約 -->
<target-concept-constraint ref="meta:人間"/>
<!-- アスペクト概念制約-->
<aspect-concept-constraint ref="meta:組織"/>
<!-- ロールメタ定義 -->
<role-meta-def id="cam:メンバーRM" rel="meta:member-of"/>
<role-meta-def id="cam:リーダーRM" extend="cam:メンバーRM"
rel="meta:leader-of" cardinality="1"/>
<role-meta-def id="cam:役付メンバーRM"
extend="cam:メンバーRM" rel="meta:role-member-of"/>
.....
</cam>

```

図4 CAM定義例(組織-人間CAM)

またそれらの基本CAMを派生して定義する派生CAMとして、あらかじめ15種類用意しており、人工物の機能構造を表現する機能構造CAM、活動への参加を表現する活動参加CAM、所有関係を表現する所有-被所有CAM、状態に依存する概念を表現する状態CAM、アスペクト依存概念を補完する概念を表現する補完CAMなどがある。

3.1 CAメタ定義

図4に組織に対して人間が担う役割の一般規約を定義したメタ定義を示している。以下で図3を簡単に説明する。メタ定義の語彙はmeta及びcam名前空間に所属するものとする。

- 1) CAM-IDと上位CAM: 全てのCAメタ定義は一意に識別可能なURIをid属性の値としてもつ。また上位のCAメタ定義がある場合、そのIDをsuper属性に指定する。
- 2) 対象概念制約: CA定義に記述する対象概念をインスタンス化可能な最も典型的な概念(meta:人間)のIDをref属性に指定する。
- 3) アスペクト概念制約: CA定義に記述するアスペクト概念をインスタンス化可能な最も典型的な概念(meta:組織)のIDをref属性に指定する。
- 4) ロールメタ定義: CA定義で定義されるルール定義のメタ定義である。基本的な属性としてはロールメタ定義ID(id), 上位ロールメタ定義(extend), ロールの成立条件型がある。図ではidがcam:リーダーRMのロールメタ定義は、上位ロールメタ定義がcam:メンバーRM, ロールの成立条件型がリーダーとしての資格が与えられていること(meta:leader-of)であることが記述されている。

2.1節で説明を避けたCA定義の定義型(type)には、CAメタ定義のIDを指定することになる。また、同様に、CA定義のロール定義(role-def)中のロール型(type)には、ロールメタ定義(role-meta-def)のIDを指定する。

3.2 CAメタ定義の役割

CAMLの設計にあたっては次の2つのユーザ層を想定した。
上流CAオーサ: CAの上流階層であるCAメタ定義を定義・洗練する役割を担う。CA知識の体系的な蓄積・洗練に責任を持つ。

ベースレベルオーサ: CAの上流階層定義であるCAメタ定義を具体化したうえで、基本オントロジーに適用しオントロジーを定義するユーザ。具体的な対象のオントロジー構築を通じて経験知を蓄積する。

CAメタ定義は、CAMLにビルトインの最小限の基礎論的知見を表すと述べたが、閉じた定義とは考えていない。上流CAオーサには、コミュニティメンバのベースレベルオーサが積み上げた経験値を汎化・体系化してCAメタ定義を追加することを許している。このことが、序論で述べた、

- オントロジー構築コミュニティがオントロジー構築の経験を明示化して積み上げ
- その経験知をコミュニティ自らが汎化・体系化して表現し共有する

ことを可能にすると考えている。もちろん、CA メタ定義語彙の追加にともなって言語処理系・オントロジー構築環境の機能(新しい編集画面・整合性チェック機能など)を追加する必要がある場合は、個別にプログラムを書く必要がある。

4. CAML 処理系を中心としたオントロジー構築環境の概要

本節では、CAML 処理系を中心としたオントロジー構築環境のうち、ベースレベルオーサに対する支援について述べる。ベースレベルオーサに与えられた責務は、CA の上流階層定義である CA メタ定義を CA として具体化すること、その CA を基本オントロジーに展開し OWL オントロジーを定義することである。

4.1 CA 定義支援

オントロジー構築環境にはあらかじめ CA メタ定義が読み込まれており、その定義に対応する編集画面やメニューが生成される。ベースレベルオーサが構築環境に組み込まれた CA メタ定義から峻別したい概念に対応するものを選ぶと、定義に対応する編集画面やメニューが生成され、CA メタ定義で規定された範囲で CA を定義できるようになる。

4.2 CA 展開支援

CA を基本オントロジーに展開する場面では、ベースレベルオーサが基本オントロジーのAspect依存性を、Yarn を使って記述する。構築環境はその記述に従って自動的に OWL オントロジーを生成する。

まずベースレベルオーサによる Yarn の記述は CA 定義に従って制限され、対象概念(assign-target-concept)と参照概念(assign-aspect-concept)に指定できるのは、CA で指定された概念またはその下位概念、Aspect依存概念(role-holder)に指定できるのは対象概念の下位概念に制限される。

次に Yarn が記述されると、構築環境は OWL オントロジーを出力する。CA の記述と OWL の形式的表現は対応づけられており、オーサは OWL の論理表現を意識する必要はない。

図 2 の CA、図 3 の Yarn として CA 展開した OWL 形式のオントロジーを図 5 に示す。まず CA のロール定義の ref 属性で指定されたロールの成立条件(doctor-of, member-of)をプロパティとして出力している。

これらのプロパティを用いてプロパティ制約をクラス公理として記述する。Aspect概念(会社医務室)は部分公理(必要条件, OWL 抽象構文では partial と表現)、Aspect依存概念(産

```

ObjectProperty(doctor-of Inverse(has-doctor) super(member-of))
ObjectProperty(member-of Inverse(has-member))
Class(人間 partial
  Annotation(ca:as-TC-to ca:医療組織-人間 CA))
Class(産業医 complete 人間
  Annotation(ca:act-as ca:医師 R)
  Restriction(doctor-of someValuesFrom(会社医務室)))
Class(医務室職員 complete 人間
  Annotation(ca:act-as ca:職員 R)
  Restriction(member-of someValuesFrom(会社医務室)))
Class(会社医務室 partial 医療組織
  Annotation(ca:as-AC-to ca:医療組織-人間 CA)
  Restriction(has-doctor someValuesFrom(人間))
  Restriction(has-member someValuesFrom(人間)))
  ...
    
```

図 5 CA 展開されたオントロジー(OWL 抽象構文)

業医, 医務室職員)は完全公理(必要十分条件, complete と表現)として出力する。Aspect依存概念を完全公理として定義することで、クラスの包摂関係(産業医 is-a 医務室職員)が導出されるようになる。

Yarn によって CA と対応づけられたクラスは、独自に定義したアノテーションプロパティによってタグ付けされる。産業医は、アノテーションプロパティ ca:act-as によって、ca:医師 R を展開して定義されたAspect依存概念であることが明示される。

5. まとめ

本稿では、オントロジー構築スキルを十分に習熟していないオーサに対して形式面・内容面双方の支援を提供することを目的として、概念化Aspect記述言語 CAML を提案し、概念化Aspectがオントロジー構築において果たす役割について考察した。概念化Aspectはロール概念を認知・峻別する経験知を蓄積・共有する基礎となる表現を与えるとともに、その経験値を汎化・体系化する基盤としての役割を果たし、序論で述べたオントロジー構築の経験知を共有するプロセスの基礎となる。

紙面の都合で CAML 処理系の機能・DL 推論でできることについて述べられなかったが、それも本研究の特長のひとつである。DL 推論を用いたオントロジーの形式的な整合性検証を行う関連研究に Rector らによる[Rector 02]がある。彼らは大規模オントロジー構築経験からオントロジーのメンテナンス性を向上させることを目的に、DL 推論が適切に機能するための概念峻別と OWL 表現のガイドライン[Rector 04]を提案し、その表現を支援するウィザードを彼らのオントロジー構築環境 Protege-OWL に実装している。

一方、我々の研究では、オントロジーリテラシーが十分に高くないオーサによる概念峻別を支援することに重点をおいている。そのため CA 展開した OWL オントロジーに対する DL 推論を用いた形式的な整合性検証に加えて、基礎論的立場の知見を表現した CA メタ定義に対して CA は内容的な整合性維持機能も提供する。さらにコミュニティのオントロジー構築経験を上流 CA オーサが CA メタ定義として汎化・体系化して、コミュニティの経験知として蓄積・共有される仕組みを持つため、内容面での支援が持続的に向上することが期待される。

参考文献

- [Hozo] 法造~オントロジー構築・利用の研究サイト~ available at <http://www.hozo.jp/hozo/>
- [溝口 05] 溝口:オントロジー工学, オーム社, 2005
- [溝口 06] 溝口他:オントロジー構築入門, オーム社, 2006
- [Rector 02] Rector, A.: "Normalisation of ontology implementations: Towards modularity, reuse, maintainability", In EKAW Workshop on Ontologies for Multiagent Systems, 2002.
- [Rector 04] Rector, A., Drummond, N., Horridge, M., et al.: "OWL Pizzas: Practical Experience of Teaching OWL DL: Common Errors & Common Patterns", 14th International Conference on Knowledge Engineering and Knowledge Management (EKAW), 2004.
- [上田 07] 上田, 池田;概念化Aspectに基づくオントロジー構築, 第 21 回人工知能学会全国大会, 1D3-2, 2007