

マッピングパターンタイポロジーに関する一考察

—機能オントロジーマッピングを例として—

An Investigation into a Mapping Pattern Typology

- Using mappings between functional ontologies as examples -

瀬川 翔
Sho Segawa

笹嶋 宗彦
Munehiko Sasajima

來村 徳信
Yoshinobu Kitamura

溝口 理一郎
Riichiro Mizoguchi

大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

In ontology mapping, source concepts do not correspond to target concepts completely because concepts in ontologies are conceptualized from different viewpoints according to developers. Thus the similarities of the matching results are different from each other. For revealing these differences of the similarities, this article presents a mapping pattern typology based on an ontological analysis. It can help users of matching results to judge the mapping whether they regard as successful or not.

1. はじめに

近年、膨大に存在する Web 上での知識や、当事者にとって暗黙的であるために共有が困難であるものづくりや医療現場における知識を、共有・再利用することを支援するために、様々な分野でオントロジーやシソーラスなどの基盤的知識体系の構築が行われている。それに伴い、同じドメインや内容に関するオントロジーが複数存在するようになり、オントロジー間の対応関係を作成し、相互運用を可能にするためのオントロジーマッピングに関する研究[市瀬 07]が盛んに行われている。

オントロジーマッピング^{*1}では、マップするソース側の語彙と、もっとも意味的に類似していると考えられる語彙をターゲット側から探し対応づける。オントロジーの各語彙にはラベルに現れる表層的な意味だけではなく、オントロジー構築者が語彙にこめた意味情報が含まれている。このため、ソース語彙にこめられた深い意味情報と完全に一致する語彙が、ターゲット側のオントロジーに存在するとは限らない。したがって、あるオントロジー A の概念 A_1, A_2, \dots, A_i とオントロジー B の概念 B_1, B_2, \dots, B_j と間に作成された n 個のマッピング関係 M_1, M_2, \dots, M_n には、マッピングごとに意味の類似度のレベルの違い(マッピングの類似度と呼ぶ)が存在すると考えられる。マッピング結果を利用する各オントロジーの利用者にとって、作成された n 個のマッピングの間で、どのようにマッピングの類似度が違うのかを知ることは、相互運用する上で非常に重要である。また、マッピングの間に暗黙的に存在するマッピングの類似度の違いを明らかにすることで、マッピングされた対応付け結果の中で、どれを成功・不成功と決定するかの判断基準を明確にすることができる。

本研究ではこのようなマッピングに潜む類似度の違いを明らかにするために、オントロジーのマッピングにどのようなパターンが存在するかについて考察・検証を行う。オントロジーマッピングパターンに関する研究は、ソース語彙が 1 つのターゲット語彙に対応付けられる 1 対 1 なのか、複数の語彙に対応付けられる 1 対 N なのかといった数量的、またはロジカルな語彙の関係を示す表層的なパターンに関しては、いくつか存在する[François 08]。一方で、このような語彙レベル、数量レベルのマッピングパターンでは、同じ 1 対 1 のマッピングの中にどのような類似度の違いが存在するかは暗黙的である[Ondrej 08]。

そこで、本研究では、そのような表層的な語彙レベルのマッ

連絡先: 瀬川翔, 大阪大学産業科学研究所 知識システム研究分野, 〒567-0047 大阪府茨木市美穂ヶ丘 8-1, Tel: 06-6879-8416, e-mail: segawa@ei.sanken.osaka-u.ac.jp

ングパターンに加えて、オントロジーの概念階層の各語彙がどのような視点から分類されているのか(以下、“分類視点”と呼ぶ)に着目し、マップされた語彙の分類視点の一致の仕方という観点から、深い意味レベルでのマッピングパターンを明らかにする。考察は、設計分野における機能概念に関する語彙体系である FBRL(Function & Behavior Representation Language)語彙体系^{*2}と Reconciled Functional basis(以下 FB^{*3})とのマッピング[大久保 07, 瀬川 08]を例として、ボトムアップに行った。さらに、作成したマッピングパターンを他のオントロジーマッピングへ適用することで、マッピングパターンの妥当性についても検討した。

2. マッピングパターンタイポロジー

2.1 3種類のマッピングパターン

本研究では、図 2 に示すように、マッピングするオントロジーを概念階層とその各語彙の分類視点から構成されるものとして捉え、次の 3 種類のマッピングパターンを考える。

- ① 表層マッピングパターン(S)
- ② 深層シングルパターン(DS)
- ③ 深層連結パターン(DC)

①の表層マッピングパターンは、マップされた語彙が表層レベルでどのように対応付けられているかのパターンである。例え

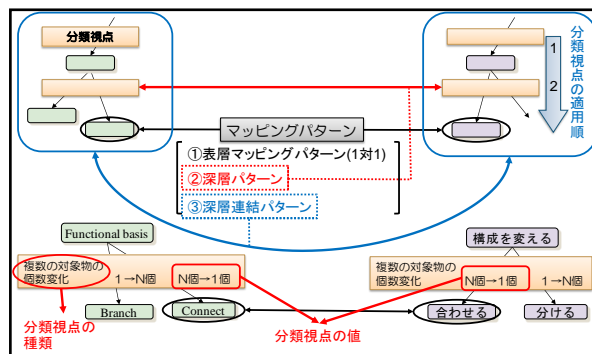


図 1 3種類のマッピングパターンと分類視点

*1 本稿では、マッピングのソースとターゲットにおいて、ソース側には何も情報を加えずにターゲットにマップすることを前提とする。これにより、マッピングは非対称な関係となる。

*2 筆者らの研究室で開発が進められ、産業界での利用実績をもつ。明示的・論理的な定義を持ち、ベース機能は 89 語彙である。

*3 米国標準技術局のプロジェクトの一環として開発された。自然言語で定義された、52 の語彙から構成される。

ば、ソース語彙が1つのターゲット語彙に1対1で対応付くか、複数の語彙のANDとして対応付けられているのか等を表す。

一方、②、③の深層シングルパターン、深層連結パターンはマップされた語彙同士の分類視点がどのように一致しているかに注目し、表層パターンでは暗黙的であるマッピングの類似度の違いを明らかにする。②深層シングルパターンは、個々の分類視点どうしを比較した場合のパターンであり、③はマップされた語彙を構成する複数の分類視点を組み合わせた場合のパターンである。

なお、これらのマッピングパターンは、オントロジーの上位レベルで対象概念の捉え方が同じことが前提となっている。筆者らが行った機能オントロジーマッピングでは、機能に関する参照オントロジー[來村 05]に基づき、機能の捉え方に関して複数の上位のオントロジカルな違いを明らかにしてマッピングを行っている[大久保 07, 瀬川 08]。このような、マップするオントロジーの語彙同士が上位のレベルで一致しているのかそうでないかの違いに関するマッピングパターン(U)は、本稿では割愛する。

2.2 分類視点の種類と値

本研究における分類視点とは、オントロジーの概念階層の各語彙がどのように分類されているのかということの根拠を表すものである。分類視点は「種類」と「値」を持つ。「種類」は、対象概念のどのような性質に注目して分類しているのかを示す。その性質がどのような「値」をとるときに各下位概念に分類されるかを表し、また、分類する性質をどのような詳細度で分類しているかを示すのが、分類視点の「値」である(図1下部参照)。

2.3 表層マッピングパターン(S)

表層マッピングパターン(S)は、マッピングのソース語彙が表層レベルとしてどのようにターゲットのオントロジーに対応付けられているかに注目し、図3のS1~S7の7パターンに分類される。図2から、異なるマッピングパターンでは類似度の違いが存在しうることが分かる。例えば、パターンS1, S2は1つのソース語彙からみれば同じように1つのターゲット語彙にマップされているが、S1は1つのソース語彙のみが1つのターゲット語彙に対応付けられているのに対して、S2は複数の語彙が同じ1つのターゲット語彙にマップされているN対1となっている。このことから、異なる表層マッピングパターンに属するマッピングどうしの類似度は異なる場合が存在することが分かる。

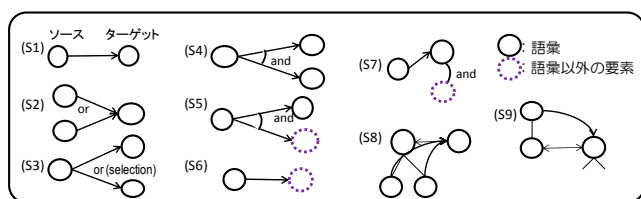


図2 表層マッピングパターン一覽

2.4 深層シングルパターン(DS)

表層マッピングパターン(S)で暗黙的なマッピングの類似度の違いを明らかにするのが、深層パターン DS, DC である。マップされた語彙 A_1 — B_1 の一段上位の語彙 Upper A_1 , Upper B_1 もマップされている場合、マップされた語彙 A_1 , B_1 を分類する分類視点とどのように一致しているかにマッピングの類似度の違いが表れると考えられる。このような、マップされた語彙同士を分類する分類視点の種類と値の個々がどのように一致しているかを考慮したパターンが深層シングルパターン DS である。深層シングルパターン DS は、次の5パターンに分類される。

- (DS1) 分類視点の種類と値が一致
- (DS2) 分類視点の種類は一致するが、値の分け方が異なる
- (DS3) 分類視点の種類は一致しないが値の相関性が高い
- (DS4) 表層パターン S5 の場合
- (DS5) 分類視点の種類も値も一致せず、相関性もない

DS1~DS4 は、マップされた語彙を分類する分類視点どうしの種類や値が何らかの形で一致する、または相関性が高いパターンである。それに対して、個々の分類視点の種類も値も一致せず相関性もないパターンが DS5 である。以下で順に各パターンについて説明する。

DS1 は、マップされた語彙同士の分類視点の種類と個々の値が一致する最も一致度が高いパターンである。ここでの例題である機能概念を分類する場合で言えば、例えば、「機能が作用対象とする(装置を流れる)対象物の数が装置の入出力の前後でどのように変化したか」ということに基づいて機能概念を分類する分類視点(「入出力の対象物の個数変化」と呼ぶ)は、FB と FBRL 語彙体系の両方に存在する。さらに、その値はどちらの語彙体系における分類においても、「N 個から 1 個へ変化」or 「1 個から N 個へ変化」という 2 値のいずれかを取り、互いの値も一致する。したがって、これらの分類視点によって分類された、「Connect」という FB 語彙と「合わせる」という FBRL 語彙の間のマッピングは DS1 という深層シングルパターンに分類される(表層パターンは S1 に属する)。

次に、DS2 は、種類は一致するがその値の詳細度が異なるパターンである。例として、装置から複数の対象物が出力される場合に「それらの対象物はどうな種類関係にあるのか」に基づいて分類する分類視点「複数対象物間の種類関係」を用いて説明する。FB, FBRL 語彙体系のどちらにもこの分類視点の種類は存在する。FB の分類視点では種類関係を「均質関係」or 「非均質関係」という 2 値に分けている。それに対して、FBRL 語彙体系は「均質関係」or 「異部分関係」or 「異成分関係」という 3 値に分けている。これは、FBRL 語彙体系では、複数対象物の種類関係を「非均質関係」から、さらに詳細な 2 つの値に分類しているためである。したがって、均質関係にある対象物の一つを取りだす「Remove」という FB 語彙から「異部分の一つをとる」or 「異成分の一つをとる」という FBRL 語彙へのマッピングは DS2 という深層シングルパターンに分類される(表層パターンは S2 に属する)。

DS3 は、分類視点の種類は一致しないが、値どうしの相関性が高いパターンである。この理由として次の3つが挙げられる。

- (1) 本質的に違う性質に基づいた分類
インスタンス集合の相関性は高いが、その理由を述べることができない(もしくは本当に偶然である)。
- (2) 本質的に同じ性質を、それと密接に関係を持つ異なる観点から分類した場合
相関性が高い理由を説明でき、本質的に同じ性質に関する分類であることが特徴である。例として、「収入の高さ」と「家の大きさ」という 2 つの種類異なる分類視点を考える。これらの分類視点は、種類は一致しないが本質的に同じ「どれくらいお金持ちか」という性質を、それと密接に関係のある「収入」と「お金」という異なる観点から分類している。このため、2 つの分類視点の値、「収入が高い」と「家が大きい」とは互いに相関性が高くなる。
- (3) (1)と(2)の中間的なパターン。

本質的に違う性質に関する分類ではあるが、分類に用いた性質の間になんらかの依存関係があり、必要条件的になっている。そのため、相関性が高い理由を説明できる。FB の複数の対象物を互いに識別(特定)可能かどうかを表す「識別可能

関係」の値“識別不可能関係”と FBRL 語彙体系の「種類関係」の値“均質関係”の一致度パターンがこれに当たる。複数の対象物が互いを識別できない関係にあるならば、対象物が“均質関係”であることが必要である。ものを識別できるかどうかの関係は、ものに対する行為(利用)に関連しているのに対して、種類関係はものが持つ性質を表しているため、本質的に異なっている。この例では、ものの性質がその利用法に基づく分類の必要条件的なものになっている。

次に、表層パターン S5 にマッピングが属する場合の深層シングルパターンが、DS4 である。この場合、表層パターン S5 は、概念階層の分類視点の値が、分類視点以外の概念と一致しマッピングされていることを表している。FB の分類視点の種類「機能タイプ」は、FBRL 語彙体系では、機能語彙体系とは別に機能タイプとして概念化されている。これは、FB は単一の語彙体系のみであるのに対して、FBRL は語彙体系と機能タイプ、メタ機能などの体系からなるオントロジーであることに起因する。

2.5 深層連結パターン(DC)

深層シングルパターンは、マップされた語彙どうしの一段上の分類視点に注目した場合のマッピングパターンであった。これを、拡張しマップされた語彙同士の意味を構成する複数の分類視点 n 個の組み合わせまで考慮したものが深層連結パターン DC である。マップされた語彙同士の意味を構成する複数の分類視点を組み合わせたときの順序や構造に注目すると、DC は次の 7 パターンに分類される(図 3 参照)。深層連結パターンにおいては、組み合わせとしてどのように一致しているのかに注目し、個々の分類視点の一致度は前述の深層シングルパターン DS1~DS4 の 4 パターンにカテゴライズされていれば、全て一致しているものとして扱う。なお、この深層連結パターンは、n 個の分類視点の組み合わせを想定して考察している。

- (DC1) 複数の分類視点と同じ順序で適用され、一致する
- (DC2) 複数の分類視点異なる順序で適用され、一致する
- (DC3) 同じ適用順序で一致するが、互いの概念階層の間に粒度異なる分類視点が存在する
- (DC4) 異なる適用順序で一致し、互いの概念階層の間に粒度異なる分類視点が存在
- (DC5) マップされた語彙の直上の分類視点は一致しないが、その一段上位と下位の分類視点と語彙は一致
- (DC6) 直上直下の分類視点は一致しない。分類粒度異なる分類視点によって分類されている
- (DC7) 個々の分類視点は一致しないが、複数の分類視点の組み合わせで表現される意味が一致する

以下では、これらの 7 パターンについて順に説明する。

DC1 は、マップされた語彙を分類する N 個の分類視点どうしが一致し、かつそれらが同じ順序で適用されているパターンである。したがって、この DC1 はマッピングを行った語彙どうしが階層的に最も類似しているパターンである。

DC2 は、マップされた語彙どうしの分類視点が一致するが、それらが互いに異なる順序で適用されているパターンである。

DC3, 4 は DC1, 2 の途中に、一方に相手にはない余分な分類視点が含まれているパターンである。

DC5 は、マップされた語彙 A, B を直接分類する分類視点どうしは一致しない。しかし、A, B の直上の語彙 Upper-A, Upper-B と直下の語彙 Lower-A, Lower-B の分類視点が一致しそれぞれマッピングされており、その階層を保存するという観点から、中間の語彙である A, B もマッピングされるパターンである。

DC6 は、語彙 A の上位(下位)の語彙 Upper(Lower)-A が語彙 B とマッピングされていて、語彙 B を更に分類する分類視点が B

に存在しないため B の下位(上位)概念が無く、語彙 A に対応する語彙が無い。そのため、語彙 A もその上位(下位)である Upper(Lower)-A と対応する B にマッピングしているパターンである。

DC7 は、マップされた語彙を分類する個々の分類視点は一致しないが、それらが組み合わせとして表現される意味が一致するパターンである。

2.6 マッピングパターンとマッピングの類似度の関係性

本節では、前節までで述べた各マッピングパターンにおいて、どのようにマッピングの類似度が異なるかについて述べる。

(1) 深層シングルパターンとマッピングの類似度

分類視点の種類と値が共に一致する場合の DS1 を基準として、他のパターンの類似度について考察する。

分類視点の種類は一致するが値の詳細度が異なる DS2 は、荒い粒度から細粒度の値へのマッピングでは 1 to N で対応付けることが可能であり、DS1 と同程度の類似度を持つ。一方、細粒度から荒い粒度の値へのマッピングでは、マップする語彙に余分な情報が含意されているため、DS1 よりも類似度は低くなる。

種類も値も一致しない DS3 の場合、相関性の根拠を表す(1)~(3)のそれぞれの場合によってマッピングの類似度が異なる。偶然一致している根拠(1)の場合には、相関性がどのくらい高いかが完全に分類視点ごとに依存するため、それらのマッピングの類似度を DS1 と比較することはできない。根拠(2)の場合には、本質的には同じ性質を密接に関係している観点から分類しているため、それによって分類される語彙どうしの類似度も非常に高く DS1 と同程度と言える。それに対して、根拠(3)の場合には、本質的には違う性質を分析しており依存関係による必要条件的に一致しているため、DS1 よりもマッピングの類似度が低いと考える。

DS4 は同じ性質に関する分類であり、概念化の手法が異なる。したがって、マッピングの類似度は DS1 と同程度であると言える。

DS5 のマッピングの類似度は明らかに DS1 よりも低い。

(2) 深層連結パターンとマッピングの類似度

次に、深層連結パターンごとの類似度の違いについて述べる。DC1 をマッピングの類似度の基準とすると、適用順序異なるが分類視点は全て一致する DC2 は DC1 と同程度のマッピングの類似度であると考えられる。

それに対して、DC3, 4 は分類視点の組み合わせ粒度異なるため、マッピングの類似度は DC1 よりも低い。

DC5, 6 は、マップされた語彙どうしを直接分類する分類視点が一致しておらず、DC1 よりもマッピングの類似度は低い。

DC7 は、組み合わせた分類視点が表現する意味の相関性が高く DS3 と類似するパターンである。そのため、相関性の高さの根拠によって類似度が異なる。

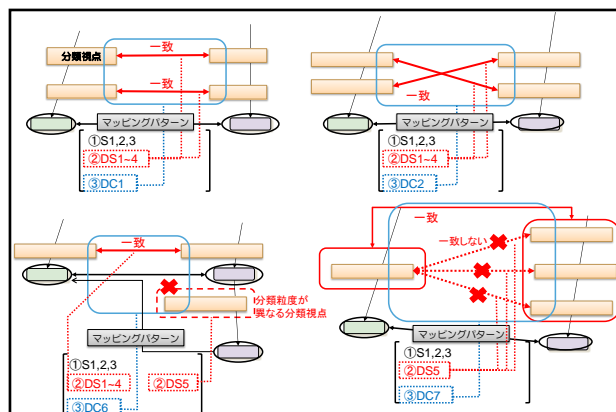


図 3 深層連結パターンの例(DC1, 2, 6, 7)

3. マッピングパターンタイポロジーの利用

3.1 機能オントロジーマッピングへの適用

前節で述べたマッピングパターンを、FB から FBRL 語彙体系へのマッピングに適用した(表 1)*4。この表から分かるように、同じ表層パターンの中にも深層パターンの異なるマッピングが複数存在する。表中の表層マッピングパターン S1 (1 対 1)には、異なる深層連結パターン DC1,2,3,6 が存在し、DC 中の深層シングルパターン DS と併せると、計 8 つのマッピングパターンが含まれている。この分析により、表層マッピングパターン S1 にも今まで暗黙的であった異なるマッピングパターン、類似度が存在することが明らかになった。これにより、なぜマッピングできたのかということに関する根拠と、類似度の違いがより明確になり、マッピングの利用者にとってマッピングの成功・不成功の基準を判断する指標として有効であると言える。

例えば、マッピング結果の利用者が改良設計などの問題解決に機能概念定義を利用する場合は、比較的厳密な対応付け結果が必要となる。この場合、2.6 節において述べたマッピングパターンによるマッピングの類似度の違いにおいて、DC1 かつ DS1 と同程度と考えられるマッピングのみを成功とみなすことが考えられる。このような成功判断基準に基づくと、表 1 の赤色部分が成功としてみなされる。このように、マッピングパターンにより今まで暗黙的であったマッピングの類似度の違いを明示化することで、マッピング結果を利用する際の利便性を高めており、利用者の相互運用性の向上にいつそう貢献できると考える。

3.2 食器オントロジーマッピングへの適用

現在、マッピングパターンの網羅性、汎用性を検証するため、機能オントロジー以外のオントロジーマッピングへの適用を行っている。本稿では、その中で 2 つの食器オントロジー A, B[山田, 05]のマッピングへの適用について述べる。上記の文献で紹介されている複数の食器オントロジーの分類視点を明示化し、マッピングを行った。そのもとで、マップされた語彙のマッピングパターンを分析し、前章で列挙したマッピングパターンがこのマッピングを網羅できるかどうかの確認を行った。図 4 に、食器オントロジー A, B においてオントロジー B から A へのマッピングの一部を示す。以下の例では、表層パターンはすべて S1 である。

まず、深層シングルパターンについて述べる。食器オントロジー中の分類視点の種類「食器の形状」は、A, B ともに存在し、値の詳細度が異なる。したがって、食器オントロジー B の「箸状食器」という語彙から食器オントロジー A の「箸」という語彙へのマッピングは DS2 に分類される。一方、食器オントロジー A の食器の用途に関する分類視点の「食器用途」の値「大型用」は、食器オントロジー B の分類視点「食器の大きさ」の値「大」と種類や値は一致しないが、相関性が高いと考えられる。これは、相関性が高い根拠(3)にあたる。食器の「用途」「大きさ」は、たがいに本質的に違う性質に関する分類ではあるが、大型用の用途に利用するためには食器が大きくなければならない、という依存関係があり、必要条件的になっているためである。このことから、食器オントロジー A の「平板食器」という語彙から食器オントロジー B の「皿」へのマッピングは DS3 に分類される。

次に、深層連結パターンについて述べる。食器オントロジー A の語彙「箸」と B の語彙「箸状食器」のマッピングに注目すると、この 2 つの概念の直上の分類視点はどちらも「食器の形状」の値「箸状」で一致している。しかし、食器オントロジー B の概念「箸状食器」の上位には、A の概念「箸」の上位にはない分類視点「食器の用途」の分類視点が存在している。したがって、この食器オントロジー B の「箸」から A の「箸状食器」へのマッピング

表 1 マッピングパターン分析(FB→FBRL 語彙体系)

表層マッピングパターン(S)	深層連結パターン(DC)	深層シングルパターン(DS)	FBの語彙数	FBの全語彙に対する比率
同一機能カテゴリ間				
1対1			33	63%
(S1) 1対1	DC1	DS1	20	37%
		DS3	7	19%
	DC2	DS4	1	2%
		DS5	1	2%
	DC3	DS1	1	2%
	DC4	DS3	1	2%
	DC6	DS1	3	6%
		DS3	3	6%
(S2) N(OR)対1	DC1	DS4	2	4%
(S3) 1対N (OR, 選択)	DC1	DS2	3	8%
	DC2	DS5	1	6%
	DC7	DS3	2	6%
(S5) 機能タイプ	DC1	DS4	6	12%

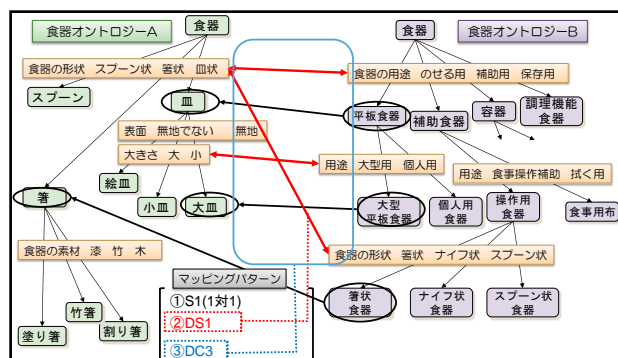


図 4 食器オントロジーマッピング(B to A, 部分)

は、同じ適用順序で一致するが、互いの語彙階層に粒度の異なる分類視点が存在する DC3 に属することが分かる。

このようにして食器オントロジーマッピングの分析を行った結果、マッピング結果を本稿マッピングパターンで全て説明することができた。このことは、マッピングパターンの妥当性、汎用性を示唆する 1 つの結果であると考えられる。

4. 今後の展望

本稿で提案した 3 種類のマッピングパターンを、今後他のオントロジーのマッピングへと適用し、その妥当性、汎用性の検証をいつそう進めていきたい。

参考文献

[市瀬 07] 市瀬龍太郎, 情報の意味的な統合とオントロジー写像, 人工知能学会論文誌, 22(6), pp. 819-825, 2007
 [來村 02] 來村徳信, 他: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(1), pp. 61-72, 2002
 [Hirtz 02] Hirtz, J., et al., "A Functional Basis for Engineering Design: Reconciling and Evolving Previous Efforts", Research in Engineering Design 13(2), pp. 65-82, 2002
 [大久保 07] 大久保公則, 他:参照オントロジーに基づいた機能オントロジーマッピング, 第 21 回人工知能学会全国大会, 2B3-2, 2007.
 [瀬川 08] 瀬川翔, 他:機能オントロジーマッピングのガイドラインに関する一考察, 第 22 回人工知能学会全国大会, 3J2-4, 2008.
 [山田 05] 山田淳, 他:博物館情報の横断検索におけるオントロジー利用の試み, 画像電子学会第 3 回画像ミュージアム研究会, 2005-03.
 [François 08] François, S., et al., "Correspondence Patterns for Ontology Alignment" Proceedings of the 16th international conference on Knowledge Engineering, pp. 83-92
 [Ondrej 08] Ondrej S., et al., "Ontological Structures through Name Pattern Tracking." Proceedings of the 16th international conference on Knowledge Engineering, pp. 213-228

*4 ここでは、2.1 節の最後で述べたように、上位カテゴリのレベルで同じ種類に分類された語彙どうしのマッピングについて議論する。