

# 実践的オントロジー構築技法－臨床医学オントロジーを例として－

## Practical Techniques for Ontology Building – A Case Study on a Clinical Medical Ontology –

古崎 晃司<sup>\*1</sup> 国府 裕子<sup>\*1</sup> 今井 健<sup>\*2</sup> 大江 和彦<sup>\*2</sup> 溝口 理一郎<sup>\*1</sup>  
 Kouji KOZAKI Hiroko KOU Takeshi IMAI Kazuhiko OHE Riichiro MIZOGUCHI

<sup>\*1</sup> 大阪大学産業科学研究所  
 The Institute of Scientific and Industrial Research (ISIR),  
 Osaka University

<sup>\*2</sup> 東京大学大学院医学系研究科  
 Department of Medical Informatics,  
 Graduate School of Medicine, the University of Tokyo

Recently, ontology attracts attention as a fundamental technology for knowledge system, and many ontologies have been developed in various domains such as bioinformatics, medicine, engineering and so on. The authors also are involved with ontology developments in several domains. This article discusses some way of thinking to develop well organized ontology through a case study on a development of a clinical medical ontology.

### 1. はじめに

近年、計算機システムを単なるデータ格納や計算のための装置ではなく、高度な意味処理が可能な知識基盤としての活用を進める動きが活性化している。その実現に向けた技術が多く研究されている中、知識を体系化し、知識処理の中核技術を提供するオントロジーへの注目が高まっている。バイオ、医療、工学など、様々な領域でオントロジー構築が進められるとともに、オントロジーを利用した情報システムの開発も多く見られるようになった。オントロジー構築には対象領域に関する知識のみならず、適切な概念化を行うためのオントロジー工学的な考察や様々な技法が必要とされる。しかし、オントロジーの構築は未だ発展途上にあり、完成された方法論は存在しない。本研究では、3年間のプロジェクト研究における臨床医学オントロジー構築の実践を通して得られた知見を、他領域にも適用可能なオントロジー構築技法として一般化することを目指す。その結果は、オントロジー構築ガイドラインの一例として、より良いオントロジー構築に貢献することが期待される。

### 2. オントロジー構築の基本思想

オントロジー構築において最も重要な点は、対象世界の本質を捉え、適切な概念化を行うことに尽きる。何をもって適切な概念化とするかを一概に定めることは難しいが、筆者らが様々な対象領域(ドメイン)において実践してきたオントロジー構築における事例を通して得られた、どのような領域のオントロジー構築の際にも共通すると思われる基本思想は、「対象世界をどのように捉えたか(概念化したか)を明らかにすることによって、諸々の概念の共通性と相違点を明確にすること」を目指した深い考察にある。

例えば、is-a 関係に基づく概念階層の構築は、オントロジー構築の要となり、is-a 階層には、そのオントロジーの対象世界にどのような概念が存在する(とみなす)か、その概念間はどうのように分類されるか、といった最も基本的な要素が現れる。しかし、オントロジーにおける is-a 階層は単なる分類階層ではなく、「どのような捉え方に基づいて is-a 階層が構築されたのか」すなわち「概念を分類する際の意味的相違点」という分類視点の明確化が重要となる。また、そのように概念間の相違点を適切に明確化するには、それらの概念に共通する性質を把握し、その共通

性を基にした同一の観点から概念を捉えることが必要となる。そのような概念が持つ本質的な性質のことを、オントロジー工学では**本質属性**と呼ぶ。概念の共通性を把握して本質概念を捉えることは、オントロジー構築において根底となる考え方である。一方、本質属性によって捉えられる概念の共通性と対極にあるのが、概念の**コンテキスト依存性**である。コンテキスト依存の概念と、コンテキスト独立の概念を適切に捉えるには、後述のロール概念に関する理論が有効である。

次節以降では、これらの基本思想のもと、臨床医学オントロジーの構築を実践することによって得られた、オントロジー構築技法について述べる。

### 3. 臨床医学オントロジー

#### 3.1 臨床医学オントロジーの概要

近年、医療情報の発達や政府による医療情報基盤構築の推進の結果、医療情報の高度な電子化が推進されてきた。現在、膨大な医療情報の電子データが蓄積され、医療サービス向上のために、その統合管理や高度な情報解析が期待されている。その情報基盤技術として不可欠なものの1つとして、厚生労働省、医療情報システム開発普及等委託研究費「医療情報システムのための医療知識基盤データベース研究開発事業」<sup>1</sup>(平成19-21年度)の一環として、国内初の本格的な臨床医学オントロジーの構築が進められている[Mizoguchi 09, 国府 08]。

#### 3.2 「法造」によるロール概念の表現

ここでは、臨床医学オントロジーの構築に用いたオントロジー構築・利用環境「法造」<sup>2</sup>における、オントロジー表現の概要を述べる。「法造」において、オントロジーは概念を表すノード、概念間の関係を表すリンクおよびスロットを用いて表される。リンクには is-a 関係を表す is-a リンクとその他の関係を表す関係リンクがあり、スロットには part-of 関係を表す p/o スロットと attribute-of 関係を表す a/o スロットがある。スロットは関係の種類(p/o または a/o)、スロットの数を規定する個数制約(Cardinality)、ロール概念、クラス制約で表される。ロール概念とは「ある“もの”が特定のコンテキストのもとで果たす役割を捉えて概念化したもの」で、ロール概念で定義される役割を担う概念はクラス制約で表される。

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:  
 06-6879-2120, E-mail: 古崎@ei.sanken.osaka-u.ac.jp

1 <http://www.m.u-tokyo.ac.jp/medinfo/medont2009proj/>  
 2 <http://www.hozo.jp>

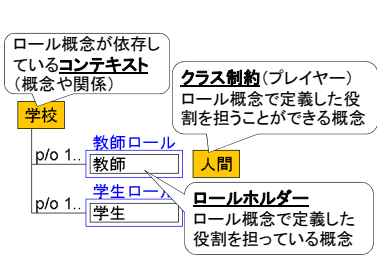


図1 「法造」のロール表現

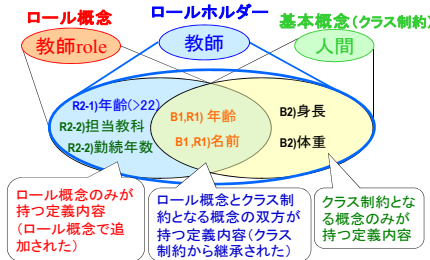


図2 教師ロールに関わる概念の関係性

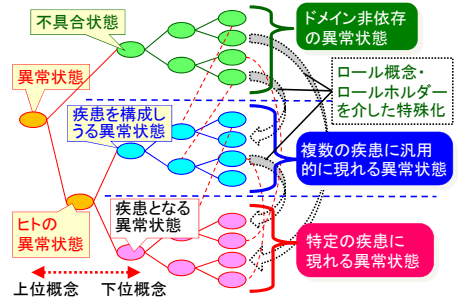


図3 疾患に関わる異常状態の関係性

現実世界の概念の多くは、様々な状況や他の概念といったコンテキストに依存して存在している。オントロジーを構築する際には、そのようなコンテキストを適切に捉えることが重要となる。オントロジー工学では、コンテキストに依存して定義される概念を「ロール」と呼び、様々な研究が行われている。法造では、ロールに対して、コンテキストに依存せず定義できる概念を「基本概念」と呼び、ロールをあるコンテキストに依存して定義される役割(ロール)を概念化した「ロール概念」と、ある概念がその役割を担った状態を表す「ロールホルダー」に詳細化して定義する枠組みを提供している(図1)。ロール概念で定義される役割を担う概念はクラス制約で表され、クラス制約は他の箇所でも定義された概念を参照している。例えば「教師」は「学校」というコンテキストに依存して定義されるロールであるのに対し、「人間」はコンテキストに依存せず定義できる基本概念である。図の法造表現では、「学校(コンテキスト)において、ある人間(クラス制約)が教師の役割(ロール概念)を担ったとき教師(ロールホルダー)と呼ばれる」という教師ロールに関わる概念の関係性が定義されている。さらに図2で示すような、これらの概念間の定義内容の包含関係を書き分ける枠組みを法造は提供している[古崎02]。このように、ロール概念をオントロジー工学の理論に基づいて扱うことができる点が法造の最大の特徴である。

## 4. 実践的オントロジー構築技法

### 4.1 概念の共通性・コンテキスト依存性の扱い

対象世界に現れる概念は、他の様々な概念に依存しており、その依存関係を明確にすることが重要となる。多くのオントロジーでみられる依存性には、対象領域による概念の違いとして現れるドメイン依存性、コンテキストの違いに依る概念的性質の違いとして現れるコンテキスト依存性などがある。ドメイン依存性は対象とするドメイン毎にオントロジーを構築することで扱うことができるが、ドメインに依存しない一般性の高い概念は、様々なドメインで共通に参照される概念として定義することが望ましい。一方、コンテキスト依存性については、3.2節で述べたロール概念を適切に用いることで、コンテキストに依存しない概念の共通な性質と、コンテキストによって異なるコンテキスト依存性を分離して定義することができる。

例として、臨床医学オントロジーにおける疾患定義に用いられる「異常状態」のオントロジーについて述べる。図3は「異常状態」の is-a 階層と各概念の相

互関係の概要を示している。異常状態は、疾患などの医療ドメインに限らず共通にみられる(ドメイン非依存な)異常を表す「不具合状態」と、医療ドメインにおけるヒトというコンテキストに依存して現れる「ヒトの異常状態」に分かれる。ヒトの異常状態はさらに、疾患の種類に依存せず共通して捉えられる「疾患を構成しうる異常状態」と、特定の疾患というコンテキストに依存して現れる「疾患となる異常状態」に分かれる。これらの異常状態の定義内容は、それぞれのコンテキスト依存性に応じたロール概念を用いて適切に書き分けられる(図4)。

例えば、「時間当たり不足状態」は「ある物質の供給量が必要量よりも少ない状態」を表す「不具合状態」として定義されて、様々なドメインにおける異常状態から参照される。その「時間当たり不足状態」が「疾患を構成しうる異常状態(複数の疾患に汎用的に現れる異常状態)」のコンテキスト下で、「不足する物質が「酸素」である」という定義の特殊化がなされ、「虚血状態」が定義される。その「虚血状態」が「心筋細胞において」現れると「心筋虚血」となり、「心筋虚血」は「心筋梗塞」という特定の疾患コンテキストにおいて「主病態の原因」となる。

このようなドメイン依存性・コンテキスト依存性を捉えた概念化は、臨床医学オントロジーにおける人体構造の概念定義にも用いられている。例えば「管状構造物」といった汎用構造物をドメイン非依存の概念と定義し、その汎用構造物を参照して「血管」や「リンパ管」といった人体コンテキストに現れる構造物を定義する。人体構造物においても、「血管」や「組織」のように様々な器官に共通に現れる人体コンテキスト下で汎用の構造と、それぞれの汎用構造が器官や臓器のコンテキストごとに特殊化される関係性を明確に分離して定義する。この際、「管状構造物」で

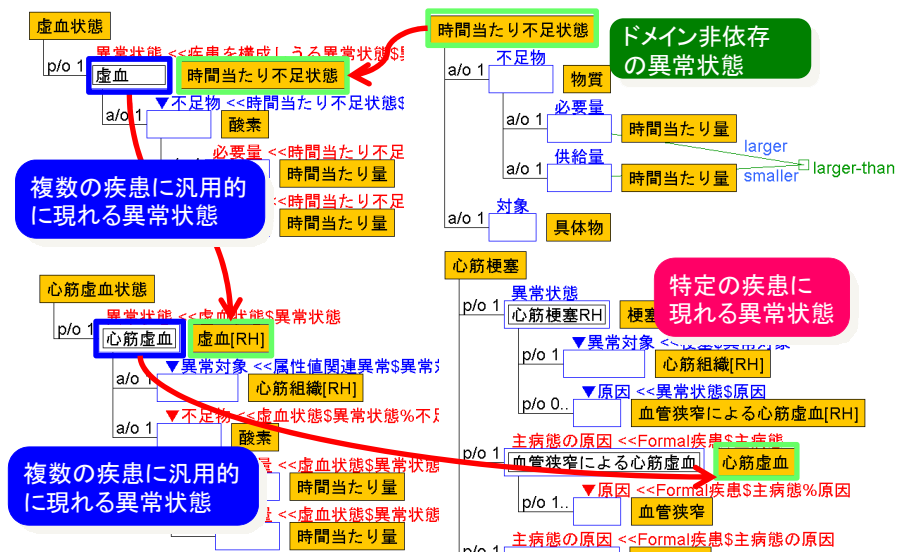


図4 臨床医学オントロジーにおける異常状態の記述例

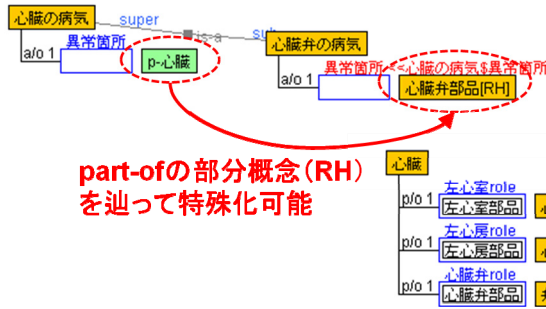


図5 p-オペレーターの使用例

あれば“断面積が小さくなって詰まることがある”，「血管」であれば“狭窄や梗塞が起こる可能性がある”というように，それぞれの汎用構造物に共通な性質を定義した上で，それらがコンテキストに依存してどのように特殊化されるかを書き分ける。

このように複雑に絡まった様々なドメインやコンテキストの依存性を解きほぐし，ドメイン・コンテキストに依存した概念と，それらから独立して共通な性質を持つ概念を分離すると共にそれら間の関係性を定義することで，明確な概念化が行える。

## 4.2 part-of 関係による部分性に基づいた is-a 関係

### (1) 疾患概念定義における is-a/part-of 関係の同時成立問題

オントロジー構築において，is-a (一般-特殊) 関係，part-of (全体-部分) 関係は最も基本的な関係として用いられ，これらの関係を適切に使い分けことが重要とされる。しかし，疾患概念を定義するにあたって，本来は part-of 関係として捉えられるべき概念間の関係が，is-a 関係としても捉えなければならない場合が見られた。例えば，「心臓」を対象箇所とする「心臓の病気」という概念と，「心臓」の部品である「心房」を対象箇所とする「心房の病気」について考える。「心房の病気」は「心臓の病気」の一種であるので，

“心房の病気 is-a 心臓の病気”

という is-a 関係が成立する。ここで，「異常箇所」を病気が持つ属性とすると，is-a 関係による上位概念から下位概念への属性継承の性質から，「心房の病気」と「心臓の病気」の異常箇所の間にも，

“心房の病気の異常箇所 is-a 心臓の病気の異常箇所”

という is-a 関係が成立しなければならない。しかし上述のように，「心房の病気」と「心臓の病気」の異常箇所のクラス制約は，それぞれ「心房」および「心臓」で，これらの概念間の関係は is-a 関係でなく part-of 関係である。これは，part-of 関係が疾患定義における異常箇所としてのコンテキストから見ると，is-a 関係として捉えられることを示唆している。

このような問題は，“(手の)指の骨折 is-a 手の骨折”のように他の疾患の is-a 関係における異常箇所のみならず，“エンジンの故障 is-a 車の故障”といった人工物，“手にさわる is-a 身体にさわる”といった行為，“神戸旅行 is-a 兵庫旅行”といった領域など，様々な概念間の関係でも同様に見られる。

### (2) p-オペレーターの導入

前項で述べた，part-of 関係にある概念間の関係が，あるコンテキストにおいては上位概念からの属性継承の性質により is-a 関係が成立すべきとなる問題を解決するため，属性のクラス制約について考察する。先にあげた「心臓の病気」を例に考えると，その下位概念となる疾患の異常箇所となりうる概念には，心房，心室，心房の壁，心房の壁を構成する組織，細胞...のように，心臓を構成する全部品が含まれると考えられる。すなわち，「心

臓の病気」の異常箇所となる概念は，心臓のみならず，心臓を構成している全部品のいずれかも含んだ概念であると言える。

このような，“全体物の部分となるもの(概念)”を一般化した概念 (“全体物自身”も部分の「部分」の1つと見なす)を表すために法造では，スロットのクラス制約記述に p-オペレーターを導入した。「心臓の病気」の例では，異常箇所のクラス制約を「心臓」ではなく，p-オペレーターを付与した「p-心臓」とすることで，心臓の病気の異常箇所となりうる概念として“心臓および心臓の部品 (part-of 関係で定義された部分概念) が含まれる”ことを表すことができる(図5)。法造は，p-オペレーターが付与された概念の part-of 関係に沿って，その全体物となる概念と部分概念の間に，概念の部分性のみに着目して抽象化した is-a 関係 (p-is-a 関係) を生成し，継承した属性(スロット)のクラス制約を特殊化する際には，p-is-a 関係に基づいた下位概念に特殊化することを認める。なお，p-is-a 関係における部分概念は，part-of スロットのクラス制約で参照している概念でなく，その全体物をコンテキストとしたロールホルダーである。また，p-is-a 関係は，その部分概念の par-of 関係に沿って，再帰的に生成される。

## 4.3 人体構造における part-of 関係の多様性問題

人体構造のような複雑な構造物のオントロジーを構築する際には，従来の part-of 関係の考え方だけでは十分に概念を捉え切れないという問題が生じた。それらの問題を解決するために，本研究では従来の part-of 関係(法造では p/o スロットで表現される)に加えて，新たに3種類の part-of 関係を導入した。以下，それぞれの part-of 関係について，導入の動機となった問題の具体例と共に述べる。これらの問題は，人体構造に限らず，構造物一般の概念定義の行方際にも共通するもので，新たに導入した part-of 関係に用いることで，様々な構造物の定義が適切に行えると思われる。

### (1) 部分分割の捉え方の多様性

複数の部分から構成される全体物は，全体物を構成している部分概念を part-of 関係を用いて表すことで定義される。概念定義の一貫性を保つためには，part-of 関係で表される複数の部分概念の間には概念的な重なりが無く，すべての部分概念を

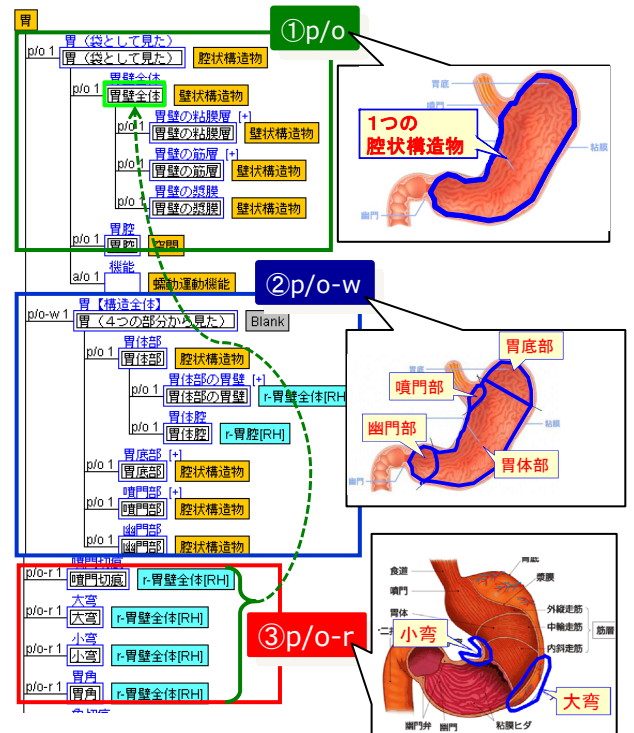


図6 p/o-w, p/o-rを用いた胃の記述例

