

Title	異常状態オントロジーとその応用
Author(s)	山縣, 友紀; 国府, 裕子; 古崎, 晃司; 今井, 健; 大江, 和彦; 溝口, 理一郎
Citation	2012年度人工知能学会全国大会 (第26回): 112R-4-3
Issue Date	2012-06-12
Type	Conference Paper
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/11213">http://hdl.handle.net/10119/11213</a>
Rights	Copyright (C) 2012 人工知能学会. 山縣 友紀, 国府 裕子, 古崎 晃司, 今井 健, 大江 和彦, 溝口 理一郎, 2012年度人工知能学会全国大会 (第26回), 2012, 112R-4-3.
Description	

## 異常状態オントロジーとその応用

## An Ontological Modeling Approach for Abnormal States

## -Interoperability of quantitative data, qualitative information, and knowledge-

山縣友紀<sup>\*1</sup> 国府裕子<sup>\*1</sup> 古崎晃司<sup>\*1</sup> 今井健<sup>\*2</sup> 大江和彦<sup>\*2</sup> 溝口理一郎<sup>\*1</sup>

Yuki YAMAGATA, Hiroko KOU, Kouji KOZAKI, Takeshi IMAI, Kazuhiko OHE, Riichiro MIZOGUCHI

<sup>\*1</sup> 大阪大学産業科学研究所<sup>\*2</sup> 東京大学大学院医学系研究科The Institute of Scientific and Industrial Research (ISIR),  
Osaka UniversityDepartment of Medical Informatics,  
Graduate School of Medicine, The University of Tokyo

Abnormality has been cross-cutting issue because of highly heterogeneous knowledge. We structured ontological model of abnormal states from generic to domain specific level. We propose unified form to describe an abnormal state as "property", and then divide it into "attribute" and "value" in qualitative form. This approach contributes to promote the interoperability and flexibility of quantitative raw data, qualitative information, and generic/abstract knowledge.

## 1. はじめに

近年、技術の進展とともに、構成部品の極小化や、それに伴う部品数の増大、多機能化が進んだ結果、専門領域がさらに細分化され、同じ専門領域(ドメイン)内でさえ知識の共有が困難な状況となり、時に深刻な事態を招く場合が生じている。一方、機械、航空、材料、医療などのドメインにおいても故障解析や診断は重要な課題であるとともに、異常状態に関する知識の記述、管理、活用は特に重要な位置を占めているという点において共通であり、その知識の共有化は信頼/安全性技術、量産技術、あるいは製品の品質向上に向けた知識基盤の構築に大きく貢献すると考えられる。

本稿では、異常状態の本質的な性質を捉えた体系化を目的とした、異常状態オントロジーの構築について、医療ドメインでの事例を中心に報告する。さらに、応用事例として、6000 強の疾患から診療科横断的な異常状態の連鎖構築の試みを紹介し、本モデルの疾患知識共有への貢献について議論する。

## 2. 異常状態の概念定義

## 2.1 異常状態定義の基本方針

異常状態に関する知識を体系化するには、対象とする異常状態概念の本質を捉え、統一した観点からの概念化が必須となる。

状態については「時間的に変動するのが当然とみなせる広義の属性が、あるときある属性値をとっていることを抽象化した概念」として捉えることができ[溝口 05]、通常は、たとえば、「空腹状態」のように、あり/なしで表される性質、すなわち特性(Property, (以下, P))に対応する概念であるといえる。

特性とは「ある属性(Attribute, (以下, A))が特定の属性値(Attribute Value, (以下, V))であるかどうか」を表す性質であり例えば、赤色は色が赤いという特定の何かには内在していない性質をさす。

ここで、医療領域における病気、工学領域における故障などを例に考えると、異常状態を議論する際には、「その状態が、異

常状態で“ある”か“ない”か」が重要視されるという点はいずれの領域において共通している。

例えば、医療分野の異常状態の例では、血糖値 250mg/dl という定量データや「血糖濃度が高い」という定性表現、「高血糖」など様々な表現がされる。しかし、これらはすべて「高血糖あり」という状態であることが本質的であるといえる。実際、糖尿病を説明する際、糖尿病の分類と診断基準による委員会報告[清野 10]では、「糖尿病はインスリン不足による慢性高血糖を主徴とし...」と糖尿病でおこる異常状態を特性で説明しているが、他の様々な異常状態においても異常状態を概念化するには、「ありなし」で表される性質、すなわち特性として捉えることが自然であり本質的といえる。

そこで本研究では、異常状態を<特性(P), 値(Property Value, (以下, Vp))>(例えば, <狭窄, あり>)と表現することとした。特性値については通常、「ありなし」の真理値をとるが、狭窄の程度が高度である場合など、その状態の程度を表す必要がある場合は、程度値をとってもよいこととした。

ここで、臨床でおこっている異常状態を概念化することを考えた場合、特性のままではその意味を十分に表現したとはいえない。例えば狭窄であれば、<断面積+小さい>という意味が埋め込まれているということ表現する必要がある。そこで、特性をさらに、対応する属性とその値のセット<属性(A), 属性値(V)>で表現する表現枠組みを導入する。詳細は 2.3 節で述べるが、これにより、異常状態のもつ本質的な性質と、実在するデータを統一的に扱うことが可能になると考える。ただし、特性のうち、本質的に特性として捉えられ、属性と属性値に分解が困難なものがある。臨床医学では対象が人体であるという性質上、正確な値の把握が困難なことが多い。例えば、変形などは特性のまま表現してもよいと考えられる。

特性の利点として、細かいパラメーター変動に影響をあまりうけないことがあげられる。また、抽象度をあげることで、診断タスクと明確に分離ができるため、疾患の原因をとらえるような状況下では、特性のままの表現が適切であると考えられる。さらに、属性と特性を分離することで、例えば「背が高い」の「高い」という属性値(V)は、英語でいう特性「tall」(P)にはならないというように、あいまいな概念を明確にすることができる。

連絡先: 山縣友紀, 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, E-mail: yamagata@ei.sanken.osaka-u.ac.jp

## 2.2 統一記述モデルの適用例とその拡張

医療ドメインでは、人体に起こる異常状態は実に多種多様であり、行為レベルから、器官、細胞レベルまで粒度もばらばらである上、関与する Agent も病原体、生体分子など様々である。そこで、本研究の提案する記述モデルが、これら異常状態を適切に表現できるか検討する。

属性は“あるモノが持つ性質”を表すので、その性質を持つ対象となるオブジェクト(Subject, (以下, O))を特定する必要がある。例えば、「胃拡張」の場合は、体積をもつものが「胃」であり、「気管支」等他の器官ではないことを示す必要がある。「高血圧」であれば、圧力を持つ対象は「血液」であることが必要である。

一方、比を表現する場合、その比をもつ対象物を特定した上で、さらに何に注目して比較を行っているかを表現しなければならない。例えば、「高血糖」の場合は、濃度は「血液」のもつ性質であるため、血圧と同様に対象物(O)は血液といえる。しかし、濃度とは、一般に、全体物に対する注目物の割合を示すため、全体物である血液に加え、注目物である「糖」も表現する必要がある。そこで、対象物(O)以外に、注目物を表す「副対象物(Sub-Object, (以下, So))」を導入することとした。

次に、「上気道ウイルス付着」では、上気道を対象とした場合、ウイルスが付着しているかどうかは、本来上気道自体がもつ性質ではない。しかし、「カーブが多い道路」において、カーブの多さを道路の属性と見なす考えと同様、臨床的にはウイルスが付着した上気道とそうでない上気道を区別する意味で、「ウイルスの付着の有無」は上気道の性質(メタ属性)属性として扱うことができる。その時、何の有無、あるいは「何の」個数であるか、この場合であれば、「ウイルスの」ということを示す必要が生じる。そこで、回数や個数の場合も、「副対象物(So)」を設け、「何の」(個数/回数)という情報を概念化することとした。

その結果、比の場合も、メタ属性についてもその対象を明確に分離しつつ、副対象物(So)を設定し、各々に必要な情報を概念化することで、柔軟な表現が可能となる。(表1)。

表1 異常状態の表現例

異常状態 (P)	特性値 (Vp)	属性 (A)	属性値 (V)	対象物 (O)	副対象物 (So)	備考
高血圧	あり	圧力	高い	血液		
冠動脈狭窄	あり	断面積	小さい	冠動脈		
悪心	あり			患者		
高血糖	あり	濃度	高い	血液	糖(グルコース)	比(全体物とその部分との関係)
上気道ウイルス付着	あり	個数	多い	上気道	ウイルス	メタ属性

## 2.3 特性属性変換: 計測データから知識の共通性抽出

各医療機関には個々の患者の膨大な診断/検査データが存在し、通常、血圧 150mmHg のように定量化されたデータとして計測される。定量データは、その対象(O)を特定した上で、属性(A)と定量値(V)を同定することができ、血圧 150mmHg であれば、対象が血液で、圧力(A)が 150mmHg という値(V)をもっているといえる。そこで、定量値にそのドメインで一般的に用いられる閾値(例: 130mmHg)を設定したうえで定性化(高い)すると、血圧 150mmHg は、血液(O) < 圧力(A), 高い(V) > という定性表現として変換され、血液を対象とし、圧力という属性が高い性質をもつ「高血圧」という特性に変換可能となる。つまり、異常状態は定量データを扱うためには、< O, A, V > で表現が必要となる。これを定性化することで < O, P, Vp > の表現に互換可能となり、異常状態としての「高血圧」あり/なしで議論することが可能になるといえる。

以上をまとめると、異常状態の本質を捉えるために、特性でまず捉えることの重要性を考察してきた。また、定量データと定性値を特性属性により相互変換可能であることを示した。そして、異常状態の表現には特性(P)、特性値(Vp)、属性(A)、属性値(V)、に加え、対象物(O)を特定する必要がある、さらに比などを表現するには、副対象物(So)を構成要素とするという表現枠組みを提案した。この表現により、相互運用可能な 3 種類の記述方法を得た。

- (1)計測データの表現として < O, A, V >
- (2)異常状態の表現として < O, P, Vp >
- (3)異常状態の特殊例として < O, P, Vp, So >

## 3. 異常状態の is-a 階層

これまで、異常状態の統一的な記述形式について述べた。これにより、様々な異常状態の共通性を把握することができる準備が整ったことになる。そこで、本章では異常状態の is-a 階層構築について述べる。

異常状態の is-a 階層を構築するにあたり、ある故障や疾患といった特定のコンテキストに依存してあらわれる異常状態の性質と、コンテキスト独立に抽象化して捉えた異常状態に見られる共通な性質を、概念として適切に扱えなければ、その世界を対象とする異常状態を必要十分に表現したとはいえない。そこで、本研究では、異常状態概念を下記のような 3 層構造として捉えることとした(図 1(a))。

- レベル1: 汎用異常状態(対象物非依存の異常状態)
- レベル2: 対象物に応じて共通に現れうる異常状態
- レベル3: 特定のコンテキスト(疾患)を構成する異常状態

以下、各レベルについて詳述する。

### 3.1 レベル1: 対象物非依存の異常状態

最上位(レベル1)の異常状態は、複数の対象物に汎用的に見られる共通の異常状態(汎用異常状態)として捉えられる。例えば、亀裂や変形、変色、機能低下等は構造物を特定せずに定義することができる(対象物非依存)、プリミティブな異常状態であるといえる。このように、極めて一般性の高い概念として、異常状態を捉えることで、機械、材料、航空、医療など複数のドメインで共通に用いることが可能である。

一般的な異常状態を考えた場合、構造に関する異常、および、発揮する機能が何らかの異常を示した状態の 2 種類が、代表的な異常状態の種類として考えられる。そこで、本研究では、異常状態を、構造関連異常、機能関連異常、その他の異常に分類することとした(図 1(b))。構造関連異常は構造に関する異常状態であり、下位概念には、変形など構造全体の形状異常、硬化など材質の異常、逆位等位置(トポロジー)異常や構造欠損/過剰など構造のアイデンティティを保ちながらもその構成が変わる構造異常を含む。これらは具体的な対象物の種類を特定せずに、汎用的に定義することができる。対象物を特定した異常状態は、これらの汎用異常状態を参照し、レベル2以降で定義される。例えば、「面積縮小」は、構造関連異常の下位のサイズ異常として定義され、様々なドメインにおける異常状態から参照される。

機能関連異常とは、機能物が発揮する機能に何らかの異常が生じた状態をいい、対象物や機能の種類を特定しない汎用異常状態としては、機能亢進、機能低下、機能停止、機能不全などが定義できる。



図1 異常状態の is-a 階層

構造、機能以外のその他の異常において、パラメーター異常は、閾値と比較して属性値が大きい状態であるパラメーター増加(大きい)異常と小さい状態のパラメーター低下(小さい)異常に分けられ、下位には、各々速度、重さ、圧力、濃度など一般性のある物理的なパラメーターを用いて、速度増加/減少、重さ増加/減少などが概念化されている。

### 3.2 レベル2: 対象物に応じて共通に現れうる異常状態

レベル2では、対象となるオブジェクト(対象物)に固有の異常状態が定義される。レベル2最上位では、複数ドメインに共通する形態的全体物を対象とした異常状態が定義される。形態的全体物とは、形態そのものにアイデンティティをもつオブジェクトであり、例えば、中空性構造物、袋状構造物などをいう。レベル1で定義された汎用的な異常状態を異常対象に沿って特殊化することで、レベル2の異常状態の is-a 階層が定義される。例えば、水道管、給油管、消化管等は、内腔と壁をもつ汎用的な構造概念「管状構造物」として捉えることができる。レベル1の「面積縮小」は、管状構造物を異常対象とした場合、管の内腔断面積が縮小した「管狭窄」という管固有の異常状態に特殊化される。

レベル2の更に下位に行くにつれて、対象となるドメインのオブジェクト(対象物)下で現れる共通の異常状態を担う概念が定義されていく。医療ドメインを例に説明すると、レベル2最上位の管状構造物は、人体固有の「人体管状構造物」コンテキストで先の「管狭窄」を特殊化した「人体管狭窄」となって定義される。さらに、それが血管で現れると「血管狭窄」となり、冠状動脈では「冠状動脈狭窄」となるように、そのドメインで必要とされる粒度まで階層記述を行う。

ここで、オントロジー構築について、下位概念をどこまで詳細にまで構築するかが課題となる。例えば、医療であれば「ポリープ形成」を場所が特定されただけの「胃ポリープ形成」状態まで概念化しておくのは理論上、冗長であると考えられる。しかし、各ドメインで、現実の状況に適用することを考慮した場合、疾患や故障では、異常状態は、隣接器官/構造物が互いに影響を及ぼし、その結果次の異常状態を引き起こすことが多いと考えられる。例えば、上述の胃ポリープは鼻ポリープとは明らかに性状も異なっており、胃ポリープは胃がんの危険性を生じ、鼻ポリープは嗅覚異常を引き起こすという点において、どの器官/構造物でおこっているかが重要となる。そこで、対象とするドメインで利用されるために必要十分な状態をモデル化するというオントロジー工学的な立場をとった場合、異常状態オントロジーの構築においては、場と取り巻く環境を詳細レベルまで考慮する必要が

あり、下位概念として、特定の器官/組織/細胞レベルの異常状態まで用意することが妥当であると考えられる。

ただし、レベル2で定義される異常状態は対象物がどこまで詳細に特定されたとしても、どのようなコンテキストで現れたものであるかというコンテキストとは独立な点において次のレベル3と概念的に分離されるべきものであり、特定の疾患に現れる異常状態とは明確に区別される。例えば、「高血圧」は「動脈硬化性症」という疾患の原因となることが知られているが、他の疾患でも汎用的に現れうる異常状態である。ここで「高血圧」が「血圧が通常の健常者よりも高い状態」であるという性質は、いずれの疾患で現れるかというコンテキストに独立して定義できる共通な性質であるが、どのような機序で現れ他にどのような影響を及ぼすかといった性質は、高血圧が現れる疾患のコンテキストに依存した性質であると言える。さらに、「高血圧」をより一般化した「圧力増加異常」を考えるならば、それは医療に限らず様々なドメインにおいて一般的におこりうる異常状態であるといえる。

### 3.3 レベル3: 特定のコンテキストを構成する異常状態

レベル3はレベル2で定義した異常状態が、ある疾患や故障という特定コンテキスト下で、それを構成する異常状態となっているものと捉えられる。例えば、冠状動脈で起こりうる異常「冠状動脈狭窄」は疾患「狭心症」を構成する異常として現れ、心筋虚血を引き起こす原因となり(図 1(c)右下)、心筋梗塞では原因、糖尿病においては結果として現れることが知られている。

このような異常状態の is-a 階層構築により、上位の抽象化レベルがあがるほど、分野を超えて共有できる知識となる。また、下位レベルでは各ドメインの専門家の細分化された知識を記述することで、固有の概念をも必要十分に表現できるという点において実用的である。このように、汎用的な異常状態から特殊化された異常状態までいずれのレベルも計算機で扱うことができることが利点である。

本階層構造によって、共通性と特殊性を明確に分離するとともに、コンテキストに応じた異常状態を表現することが可能となることから、異常状態の本質を捉えた概念化を基盤の構築はもとより、拡張性のある多様な応用へも対応できると考えられる。

## 4. 医療への応用

これまで、異常状態の本質を理解するために、異常状態の表現形式と is-a 階層構築について提案した。本章では、これらに基づいた応用について医療を例に説明する。

医療分野では、複雑で多様な疾患概念を的確に捉えたモデルが必須となる。その際、特に重要な概念が異常状態であると



いえる。疾患と異常状態の関係を考察すると、疾患では異常状態が次の異常状態を引き起こす原因となり、伝播しつつ、次の連鎖が進んでいき、重篤な影響を及ぼしていくことが見出せる。つまり、疾患は異常状態から構成され、疾患を異常状態の原因、結果の連鎖の総体と定義できる[Mizoguchi 11]

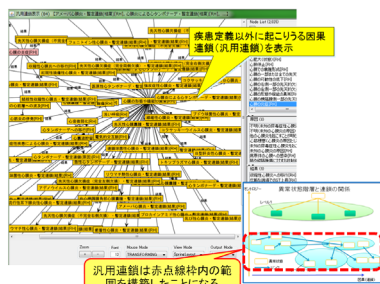
筆者らが参画している臨床医学オントロジー構築プロジェクトでは、臨床専門医が各診療領域で対象とされる個々の疾患について、異常状態ノードとその関係を表すリンクから構成される有向グラフで表現し(疾患連鎖)、異常状態の対応付けを行っている。各疾患で現れる異常状態(疾患コンテキスト依存)は、3章(4)で説明したレベル3の異常状態の原因または結果として定義される。原因、結果をたどることで、異常状態を動的につなげた連鎖(汎用連鎖)の生成も可能となる(図 2(a))。現在、12 診療科で約 6000 の疾患と 14000 を超える異常状態について連鎖の記述をすすめるとともに、2 章の異常状態表現に基づいた統一記述の洗練作業をすすめている。[大江 10]。また、これと並行して、現在レベル1の汎用的異常状態と、レベル2の構造物固有の異常状態も構築中である。先のレベル3の疾患を構成する異常状態は、レベル2のリーフに各器官固有に現れうる異常状態という知識としても蓄えられている(図 2(b))。これらは、他疾患で同様の異常状態が起こった場合、その疾患を構成する異常状態という形で再利用されることとなる。

レベル2以上は全診療科で共通して参照される異常状態であり、例えば、循環器内科で「冠状動脈狭窄」状態が記述され

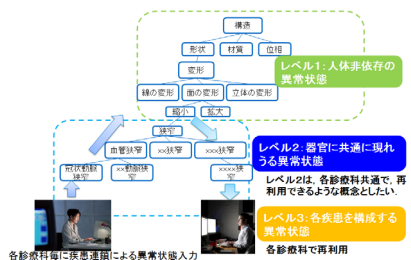
た場合、それは、レベル2の is-a 階層の最下位(リーフ)の異常状態として保存され、他の診療科、例えば、代謝内分泌科の糖尿病の高血糖が引き起こす結果としても利用が可能となる。また、消化器内科の腸閉塞の原因である「腸管狭窄」とは、同じ「管の狭窄」という上位概念が共有される。このように、各診療科(専門領域)で詳細に記述された粒度も異なる異常状態を汎用的な異常状態概念として定義し、全診療科で共有することにより、異常状態に関する知識基盤構築に貢献すると考えられる。これらの知識基盤を用いて、異常状態の連鎖(水平展開)と is-a 階層構造(垂直構造)の展開を俯瞰すると、断片的な知識でしかなかった異常状態から、全体像および疾患/故障に至るプロセスを統合的に捉えることができると考えられる。さらに将来的には、下位の異常状態同士の因果関係(連鎖)から上位階層同士の因果について推論への応用が期待できる(図 2(c))。例えば、すでに究明の進んでいる異常状態連鎖から得たシナリオから、垂直方向の関係を追跡することで上位の連鎖の共通性によって相互運用性の高いシナリオを提供し、ある疾患の未だ不明となっている中間の異常状態を推論できる可能性がある。

さて、本研究が提供する統一記述モデルは、2.3 節で述べたように特性属性変換により定量データと抽象的な知識を変換できるという利点もある。例えば、「頸動脈内径 3mm」というデータは、「頸動脈の内径が短い」と定性化され、さらに「頸動脈狭窄」と変換することができる(図 2(d))。したがって、臨床データから抽象的な疾患知識までの情報が計算機が処理できる形で統管理し、相互運用可能なシステム開発の実現にも貢献できる可能性がある。

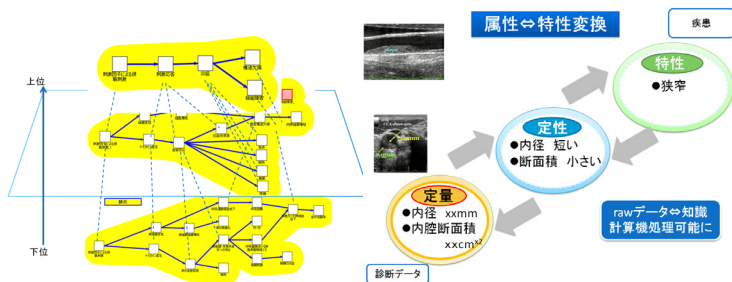
汎用連鎖の動的生成と可視化



a) 疾患における汎用連鎖



b) 疾患の異常状態 is-a 階層



c) 疾患連鎖レイヤーイメージ

d) 属性特性変換

図 2 医療分野における応用例

## 5. まとめ

本研究では、複雑で多様な異常状態概念のオントロジーを構築することによって、知識の共通性が理解できるとともに、分野横断的かつ汎用的な状態から、ある特定の分野でのコンテキスト依存状態までの体系化を可能とした。また、異常状態の特性記述と属性分解により、異常状態の統一的な記述モデルを提案した。さらに、異常状態の因果関係を連鎖で捉えることで、故障に至るプロセスを捉え、水平、垂直構造を組み合わせていけば、医療分野に限らず、異常状態の全体像をとらえ、FMEA、FTなどの故障解析にも応用できることが期待できる。

このように本研究は、異常状態の体系的知識構築と統合理解のための枠組みを提供し、分野内から分野横断的な多くの人が共有する幅広い知識の共有まで、活用による相互運用性の向上に貢献する可能性をもつと考えられる。

## 謝辞

本研究の一部は、日本学術振興会の最先端研究開発支援プログラムおよび厚生労働省医療知識基盤研究開発事業により、助成を受けたものである。

## 参考文献

[溝口 05] 溝口理一郎 オントロジー工学, オーム社, 2005  
 [清野 10] 清野裕他 日本糖尿病学会糖尿病診断基準に関する調査検討委員会, 糖尿病の分類と診断基準に関する委員会報告, 糖尿病, Vol. 53, pp450-467, 2010.  
 [大江 10] 大江, 今井: 臨床医学知識処理を目指した医療オントロジー開発, 人工知能学会誌, Vol. 25, No. 4, pp.493-500, 2010.  
 [Mizoguchi 11] R. Mizoguchi, et al., In Proc. of International Conference on Biomedical Ontology(ICBO2011), Buffalo, NY, June 26-30, 2011.