

臨床医療オントロジーの構築に関する基礎的な考察

A Fundamental Consideration toward Development of Medical Ontology

国府裕子^{*1} 周俊^{*1} 古崎晃司^{*1} 今井健^{*2} 大江和彦^{*2} 溝口理一郎^{*1}
Hiroko KOU, Jun ZHOU, Kouji KOZAKI, Takeshi IMAI, Kazuhiko OHE, Riichiro MIZOGUCHI

^{*1} 大阪大学産業科学研究所

^{*2} 東京大学大学院医学系研究科

The Institute of Scientific and Industrial Research (ISIR),
Osaka University

Department of Medical Informatics,
Graduate School of Medicine, The University of Tokyo

The current goal of our research is to build a Japanese medical ontology. It becomes a key to the successful development of various knowledge processing applications for systematization of vast medical knowledge and the synthetic understanding of them. This article discusses a fundamental consideration of medical ontology based on ontological theory. We focus on "anatomical structure of organs" and "abnormal state in the human body". On the basis of the investigation, we distinguish organ-specific concepts from those independent of any organ. And we described abnormal states of a human body by the same framework which is represented by an object, its attribute and its value(abnormal type). These discussions give a guideline to fill the gap between different medical domains through establishment of common concepts across various branches of medicine.

1. はじめに

近年、医療現場において多様な情報の電子化が進み、電子カルテのみならず、処方箋の内容や画像、数値の検査結果、画像検査、診断報告書などを含む多量の電子情報が蓄積されつつある。これにともない、蓄積されたデータの意味を捉え多目的な情報処理を行なうことに関心が払われつつある。たとえば診療現場における特徴をもった患者情報の効率的な検索や類似症例の検索、多数の患者データベースからの共通医学知識や相関ルールの抽出、医療事故防止支援のための情報提示など様々な知的処理が医療の質の向上へ貢献するものとして期待されている[大江 08]。この様な処理を可能にするためには、医療データに含まれる多彩な専門用語・語彙の標準表記とコード化だけでなく、それら用語の概念レベルの定義と概念間の意味的な関係性¹を明示した医療に関する知識や技術等の論理的な体系化および医療に関する情報や知識を蓄積・共有・創発できる医療知識基盤の構築が必要とされている。

2. 臨床医療オントロジーと疾患知識の体系化

臨床医療知識は、診断学をはじめとする臨床医学、製薬、機器、工学など多様な専門領域知識に渡っており、これらの複数領域の知識および概念が互いに深く関連性を有している。このため、領域を横断した医療基盤知識の体系化・共有化が強く求められている。このような背景のもと厚生労働省の「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」において、情報や知識を蓄積・共有・創発するための知識基盤の基盤となる臨床医療オントロジーの構築が進められている。筆者らはこの事業に従事し、臨床医療オントロジーの構築に向けて、オントロジー工学の基礎理論に基づいた考察を進めてきた。医療領域は様々な専門領域を取り込んだ膨大な知識の基盤の上に成り立つが、特に疾患に関わる知識は重要な基盤知識の1つである。そのため、現在までも様々なリソースによって医療のための語彙・用語の整理および標準化がなされてきた。そ

の代表例の一つであるSNOMED-CT [CAP 02]においては、人体構造について約 30,000、所見について約 40,000、疾患については約 80,000の語彙が定義されている。

本研究では臨床医療オントロジー構築に向けて、基本的な疾患知識の記述に必要な人体構造および疾患等の異常状態に関する概念を対象とした基礎的な考察を行った。ここでは、疾患および疾患に関連する症状・所見、および人体構造を中心とする考察を目的としたため、考察対象とする人体構造物の最小構成単位は基本的な疾患概念の記述に必要なレベルにとどめた。よって、ゲノムクスからのアプローチで扱われる遺伝子、および細胞周期のイベントごとの染色体動態異常など、遺伝子レベルの構造物の動態異常等は考察の対象外としている²。

3. 解剖学的人体構造の概念定義

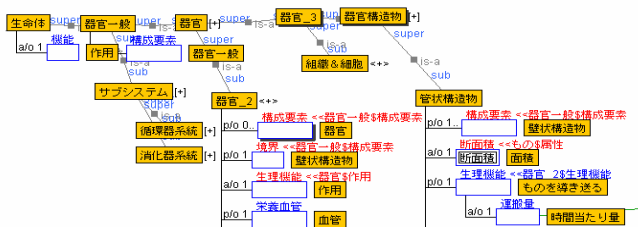
3.1 人体構造の上位概念

人体構造を大きな単位で捉えると、まずは細胞が存在し、その細胞の集合体により組織が構成され、それらの組織が複雑な構造関係を持ち、器官および様々な小器官を構成している。またある種の器官や組織はその内部に多様な位置関係で小器官を部分器官として内包する。たとえば、胃は胃腔と胃壁よりなり、胃壁は複数の細胞によって、内側から粘膜層、粘膜下層、筋層の3層構造で構成されている。その胃壁の胃底部の粘膜の胃小窩の底に、副細胞、主細胞をはじめとする多数の細胞群よりなる胃腺が胃の部分器官として、複数の層、組織内を貫通して粘膜の最下層まで横断する位置関係を呈している。

図1にこれらの人体構造物の上位概念の概要を示す。まずは上述のような人体を構成する全構成要素を表す概念を「器官一般」とし、その下位概念として心臓や肝臓といった単体で動作する「器官」(細胞、組織も含む)と、それらが複合することで一連の生体機能を担う循環器系といった「サブシステム」の2つを導入した。

連絡先: 国府裕子, 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: kou@ei.sanken.osaka-u.ac.jp

¹本稿では、同義語等の扱いに混乱をきたさないように語彙と概念を区別する。同じ意味をもつ語彙は同じ概念にマッピングされる。
²遺伝子疾患の症状および遺伝子異常状態自体は、本稿で述べる枠組みで記述できる。



※図中 p/o は part-of 関係を表す

図 1 人体構造における上位構造枠組み

さらに「器官」は心臓や胃といった「一般的な器官を表す概念」(「器官_2」と、それらの「器官を構成する構成物を表す概念」(「器官_3」)の 2 つに分けられ、心臓は前者、心臓を構成する心筋組織や細胞は後者の下位概念として定義される。

「器官一般」が細胞や組織、他の器官など様々な構成要素から構成されることは他の「器官一般」と part-of 関係で表される。「器官一般」の下位概念において、構成要素となる「器官一般」を特殊化しつつ、さらにその構成要素を再帰的に part-of 関係で表す。この再帰構造は、構成要素が細胞や組織など、その「器官一般」の末端となるまで繰り返され、これにより任意の段数から成る part-of 階層を持つ複雑な人体構造物の概念が記述可能となる。

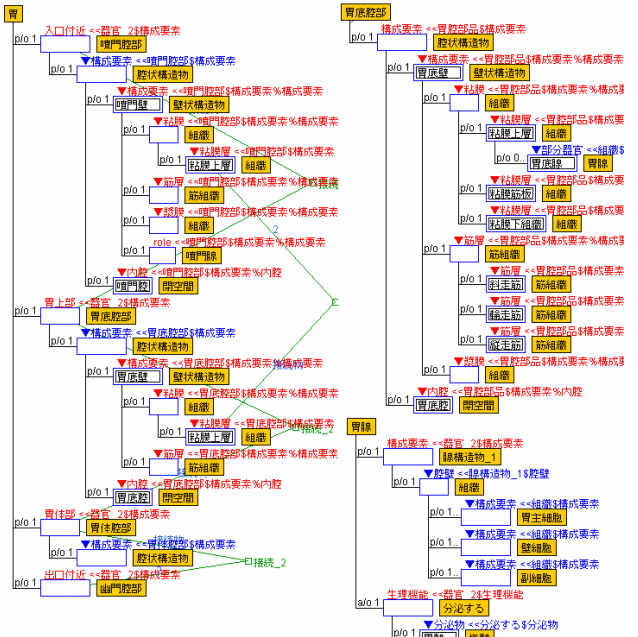


図 2 胃構造の再帰性の記述

たとえば、胃は内腔と大きく 3 層構造とされる胃壁からなり、胃壁の各層は細胞集合である組織からなる。組織は末端的構成物であると同時に構成要素として「器官一般」を持つことができるので、再び器官である「胃腺」を構成要素として記述できる。

3.2 人体構造における汎用的な構造概念

ここで、人体の器官およびその部品を持つ汎用的な構造概念について考察する。前述のように、人体における器官は多数の部分器官から構成されるが、これらの部分器官にはどの器官の部品であるかというコンテキストに依存しない汎用的な構造概念として捉えることができるものがある。たとえば、静脈や食道は共に管状、すなわち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。また、胃や心房を構成する壁は共に壁状の構造を持つ「壁状構造物」、胃や心房の中空になった部分構造は「腔状構造物」などと捉えることができる。このよう

な汎用的な構造概念を導入することで、「管状構造物」は共に詰まる(狭窄)可能性がある、「壁状構造物」は穴が空く可能性があるなど、異なる器官の疾患の共通性を扱える。たとえば、脳血管におきた狭窄も、冠動脈におきた狭窄も、管状構造物の断面積の変化としては同レベルで扱うことができる。

一方、共通の構造を持っていても、通常はどの器官の部品であるかというコンテキストに依存して異なる性質を持つ。たとえば、先に挙げた静脈と食道は同じ管状構造物であるが、それぞれのコンテキストに依存した部品としての性質も有している。具体的には、静脈の壁は大きく 3 層構造の壁を持ち、四肢や頭部などの位置により、その内皮細胞の一部が静脈弁を構成している。これに対して、食道は、壁は血管と同様に 3 層構造であるが、長軸方向に対して筋線維が斜行している 2 層構造の筋層により、蠕動運動というユニークな機能を果たしている。同じ壁でもこのような違いがあり、さらに、壁に食道噴門腺等の部分器官を再帰的に含むなどの固有構造をもつ。

このような、どの器官の一部であるかというコンテキストに依存した概念の性質を適切に捉えるために、本研究ではロール概念を用いる。ロール概念は「特定のコンテキスト下において実体が果たす役割を概念化したもの」で、この状況に依存して変化する役割を担う個物を潜在的ロールプレイヤー、個物が実際に役割を担った状態下にある時をロールホルダーと呼ぶ。[溝口 99, 溝口 05, 古崎 02, 砂川 05]。たとえば、「壁状構造物」が、胃では「胃壁」、血管では「血管の内壁」として振る舞うことをロール概念を用いて表すことができる。このような概念の峻別は、人体とその機能の関連の記述において生理的機能、構造的機能の依存先を明確にするために重要となる。

また、人体構造に関連する概念の扱いの問題点として、身体部位のような実体を持たないものが指す概念の扱いがある[荒牧 06]。たとえば鼻の内部の空間をさす鼻腔は、他の身体部位のような実体を持たないものであるが、これらを他の人体部位と同様の概念として扱ってよいのかという問題がある。この問題については「空間」という概念を導入し、鼻腔は「空間」が鼻という構造物をコンテキストとして「鼻腔」のロールを担っていると捉えることができる。これにより、「鼻腔」はそもそも実体を持たない空間であるという情報に加え、鼻という構造物の一部分を担う空間につけられた名称であることを同時に表すことができる。

3.3 人体構造の記述例

前述の枠組みと記述に基づき心臓を記述すると図 3 のようになる。心臓は二つの心房部、二つの心室部から構成され。心房部は左心房部と右心房部、心室部は左心室部と右心室部に分かれる。心房や心室を構成する壁は壁状構造物として定義されている。

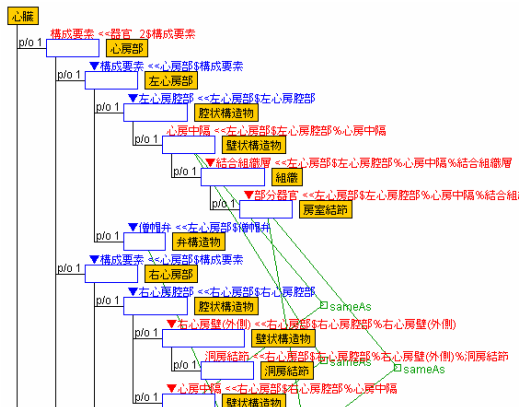


図 3 心臓の記述例(一部)

また心臓というコンテキスト下において生じた左心室と右心室の境界は心室中隔と呼ばれ、コンテキスト下における独自の構造を有するものと同時に、汎用的な壁構造を有する壁状構造物として記述されている。ここで、心臓というコンテキスト下において存在する左心房と右心房の境界部分に、例えば心房中隔などを記述する場合、心房中隔は「心臓」というコンテキスト下で境界「心臓中隔」としての役割を担うロールとして定義される。そして周囲を囲む多数の心臓壁により存在する空間に命名される左心房は、心臓壁が境界となり生み出した空間が担うロールホルダーとして定義されている。なお心房壁を構成する心筋組織の構成細胞等の性質は、それぞれのロール概念のプレイヤーである壁状構造物から継承される。

3.4 位置関係や接続関係などの構造情報の記述

人体構造を記述する際には、part-of 関係だけでは表現しきれない複雑な構造情報が多数存在する。たとえば、(1)ある器官における構成要素の一部が別の器官の構成要素となっている、(2) 任意の部品の一部に名称が当てられる、などがある。

(1)の例として、たとえば眼窩は頭蓋骨の構成要素である前頭骨の一部、および複数の骨の一部を壁としている空間を指す。このような場合、眼窩の構成要素と頭蓋骨の構成要素の構造は、単純な part-of 関係だけでは捉えきれず、異なる器官の構成要素間の内包関係もある。

(2)の例として、たとえば顔の一部分を指して頬と呼ぶ場合なども、前述と同様に器官と構成要素の間の part-of 関係に加え、構成要素間の内包関係が存在する。

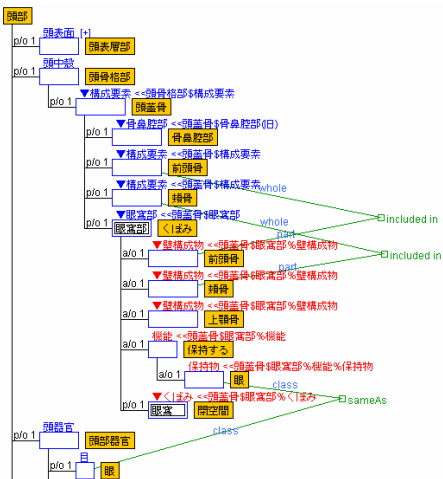


図4 位置情報や接続関係などの構造情報の記述

このような複雑な構造は、必要な関係概念 (part-of 関係以外の関係定義) を適切に導入することで表現できる。たとえば図4は、眼窩の構成要素と頭部の構成要素の間の関係を、include-in (包含) 関係を導入して明記した例を示す。このような構造情報を正確に記述することは、たとえば、眼窩を構成する頬骨の転位や眼窩の下壁骨折によって、眼窩という構造を持つ眼球を保持する物理的機能が失われ陥没が生じるといった因果関係の推論を行なう際に必要となる。

4. 異常状態としての疾患概念の記述

4.1 疾患の上位概念の概要

疾患概念を含み、疾患の原因となる異常、自覚および他覚を含む徴候や未病まで、人体で起こる様々な異常は抽象化された状態として捉えられる。そのため、疾患概念は異常概念の下位概念として定義した(図5)。異常状態の下位概念には、ドメインに依存しない汎用的な異常状態概念である「不具合状態」と、疾患概念定義に関連した「ヒトの異常状態」が定義される。そして「ヒトの異常状態」の下位概念には、「疾患となる異常状態」と「疾患を構成しうる異常状態」が定義され、医療分野で一般に疾患として捉えられている概念は「疾患となる異常状態」の下位に「Formal疾患」として定義される¹。なお本研究の考察対象は、現在はFormal疾患に限定している。

ンに依存しない汎用的な異常状態概念である「不具合状態」と、疾患概念定義に関連した「ヒトの異常状態」が定義される。そして「ヒトの異常状態」の下位概念には、「疾患となる異常状態」と「疾患を構成しうる異常状態」が定義され、医療分野で一般に疾患として捉えられている概念は「疾患となる異常状態」の下位に「Formal疾患」として定義される¹。なお本研究の考察対象は、現在はFormal疾患に限定している。

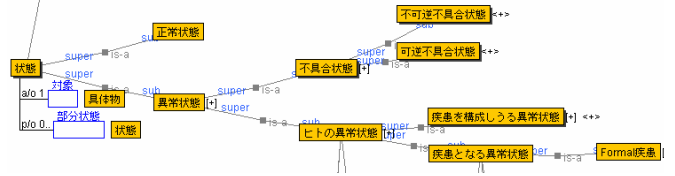


図5 疾患の上位概念

4.2 疾患概念 (Formal 疾患) の記述

疾患概念 (Formal 疾患) は、基本的に罹患者が持つ状態が異常状態であると捉え、異常状態の下位概念として定義する。異常状態は複数の異常状態より構成される。

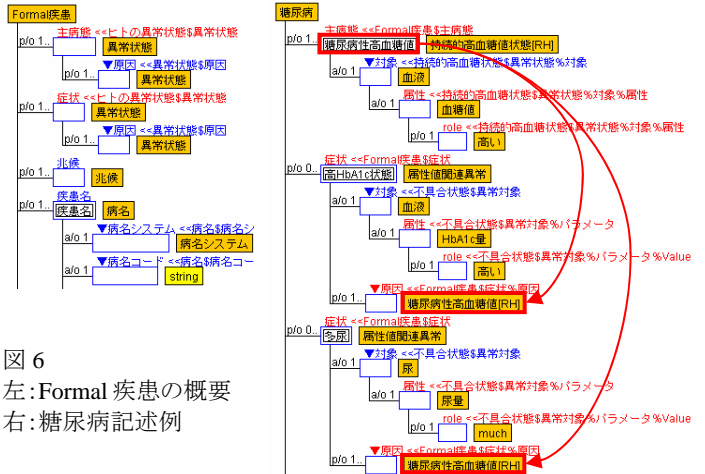


図6 左: Formal 疾患の概要 右: 糖尿病記述例

図6 左に Formal 疾患の概念定義の基本構造を示す。異常状態は複数の異常状態の組み合わせで表されるが、Formal 疾患ではそれらの異常状態を疾患は「主病態」、「症状・所見」およびそれらの原因となる異常状態に分けて記述される。主病態は対象となる疾患を医学的に定義する時に必要と思われる病態を表し、症状・所見は主病態に起因して現れる異常状態を現す。これらの異常状態の記述は「不具合状態」または「疾患を構成しうる異常状態」として定義された異常状態概念 (のロールホルダー) を参照し、すべて同型のフレームで記述される。

また、これらの異常状態の因果連鎖は、原因スロット (ロール概念) のクラス制約として他の異常状態を参照することで表される。たとえば図6右の糖尿病では、主病態である「持続的高血糖状態」が原因で「高 HbA1c 状態」や「多尿」などいくつかの症状が生じることが記述されている。このように因果連鎖は、疾患を構成する他の異常状態を参照することで表現される。

4.3 異常状態概念の記述フレーム

異常状態概念は、対象となるオブジェクトが持つパラメータの異常として捉え、「異常対象 (Object) + 異常な状態となっているパラメータ (Parameter) + 異常タイプ (Type/Value)」の組み合わせによる同型の記述形式で表す (図7)。このため、共通性が理解しやすいのと同時に、計算機での処理が容易となる。例えば、高血糖状態であれば、血液 (Object) の血糖値 (Parameter) が高

¹ 各概念のラベルは、今後の検討で変更される可能性がある。

い(Type)となる。表1にこのフレームによる異常状態の記述例を示す。

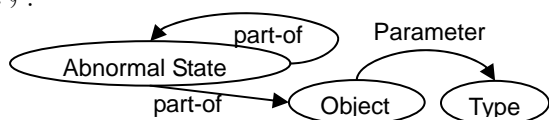


図7 異常状態の記述フレーム

4.4 異常状態概念の記述に関する検討事項

(1) 異常対象の定義

異常対象は広義の人間とし、身体箇所のみならず、人間の精神活動および生理機能、行動、動作、等を含む。よって、自傷行為といった行動における異常、視力低下や知識遅延などの機能的・能力的異常、倦怠感といった自覚的異常、および譫妄などの精神的異常、戦振など異常対象が具体的存在物でない場合も記述することができる(表1)。

表1.異常状態のフレーム記述例(仮)

Abnormal State	Object	Parameter	Type
高血糖状態	血液(身体)	血糖値(属性)	高い
血管狭窄	血管(身体)	断面積(属性)	小さい
自傷行為	行為(行動)	自傷(特性)	あり
月経不規則	月経(機能)	周期(属性)	不規則
知能遅延	知能(能力)	発達速度(属性)	遅い

たとえば、月経不規則の場合、「月経機能における月経周期というパラメータが不規則な状態にある」と記述される。ここでは、その機能を保有する機能物(生殖器といった身体箇所)の異常とはしないことに注意が必要である。機能物の異常は、「卵巣のエストロゲン分泌量の低下」のように、その原因となる異常状態として記述される¹⁾。

また、異常対象の定義において 欠如・過不足等、全身的な意味と局所的な意味の双方を有する場合がある。たとえば、脳死状態を引き起こした真の原因は脳という局所的血液量不足による酸欠だが、その背景には事故による(全身的な)大量出血などである場合がある。これらは、「脳の出血量が多い」や「全身の出血量が多い」のように、欠損物を欠損物量とした属性の異常とした上で、コンテキストに留意して、異常対象となる領域を定義する必要がある。

(2) 異常パラメータと異常状態タイプの定義

異常タイプの記述は、「過剰」、「欠乏」など定性的な属性で表現される場合と、妄想あり、振戦ありなどの異常の「あり/なし」で表現される性質がある。あり/なしで表される性質を特性と呼び、すべての属性と特性は双方向に変換ができる。このため、情報量としては同等であるが、特性は記述が容易で、医療現場の専門家が記述する際には実用的であると思われる。

(3) 汎用的な異常状態概念

疾患概念を記述する際に、疾患概念を主病態や症状で参照される異常状態は、疾患や人体というコンテキストから独立した汎用的な異常状態概念として定義した「不具合状態」「疾患を構成しうる異常」から選ばれることが多い。

不具合状態は人体、臓器、そして疾患から独立して一般化された物理現象としての異常状態を表す。例えば、管状で断面積を有する管状構造物に起こりうる一般的な物理現象としての異常状態として、攣縮、狭窄、破裂などが「不具合状態」の下位概念として定義される。また「疾患を構成しうる異常」は、虚血など、複数の疾患に共有に現れる、汎用的異常状態概念を表す。

¹⁾原因となる異常状態は、疾患や症状を定義するのに必要な階層まで追って因果連鎖が記述される。

このようにコンテキスト依存・非依存の概念を区別することで、異なる異常状態概念の共通性を明確にし、計算機上レベルで処理可能な形式で記述することが可能となる。これにより、たとえば、特定の疾患で閉じた因果連鎖のみならず、複数の疾患に出現する汎用的な異常状態を挟んで複数の疾患にまたがる因果連鎖の追跡が可能となる。

5. 総括と今後の検討課題

本研究では、広く医療情報を俯瞰した医療支援を行なうための情報基盤となる臨床医療オントロジーの構築に向けた基礎的な考察を行い、特に、疾患概念と解剖学的人体構造の記述を中心にオントロジー構築を進めた。しかし、医療知識を深く記述するには多くの課題が存在する。たとえば、医学の進歩に応じた疾患分類の変遷がある。臨床医学の知識体系は、患者に観察される症状や検査結果の組み合わせから、疾患カテゴリーを診断し、分類に応じて基本治療法を選択するという大きな枠組みから成り立っていると考えられる。しかし、症状や検査結果の組み合わせを基盤に分類されていた診断概念に、新しく開発された検査手法を適用すると、従来、同一分類とされていた疾患概念が細分化され、または異なる概念として分類されることがある。また異なる疾患と考えられていたものが、原因の解明などにより他概念と同一の分類に統合されるなどといった変化は日常的に起こる。この様な複雑な疾患分類の遷移を適切に扱うことは、1つの大きな検討課題である。

また、時間の扱いも課題の1つである。例えば脾臓は、古くなった赤血球だけを壊す機能を持つが、この機能を正確に記述するには細胞の存在時間を考慮する必要がある。また放射線治療における細胞の分裂期は放射性感受性と密接な関連性があり、セルサイクル等の循環的時間の扱いも興味深い検討課題である。

今後、知識管理・提供インフラストラクチャとして、オントロジー基礎理論に基づいた応用可能な医療知識概念の体系化を進めると共に、構築されたオントロジーの利用環境、及びアプリケーションの開発に取り組みたい。

謝辞

本研究は、厚生労働省、医療情報システム開発普及等委託研究費、「平成19年度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」の一環として行なわれた。

参考文献

[大江 08]大江和彦: 医療知識基盤としての臨床医学オントロジー, セマンティック Web コンファレンス, 2008.
 [荒牧 06] 荒牧, 今井, 梶野, 美代, 大江: 自然言語処理によるオントロジー構築支援の試みとその問題点, 言語処理学会第13回年次大会 WS 論文集, pp.51-54, 2007.
 [溝口 99] 溝口, 池田, 來村: オントロジー工学基礎論, 人工知能学会誌, Vol.14, No.6, pp1019-1032, 1999.
 [溝口 05] 溝口理一郎: オントロジー工学, オーム社, 2005.
 [古崎 02] 古崎, 來村, 池田, 溝口: 「ロール」および「関係」に関する基礎的考察に基づくオントロジー記述環境の開発, 人工知能学会論文誌, Vol.17, No.3, pp196-208, 2002.
 [砂川 05] 砂川, 古崎, 來村, 溝口: コンテキスト依存性に基づくトール概念組織化の枠組み, 人工知能学会論文誌, Vol.20, No.6, pp461-472, 2005.
 [CAP 06] College of American Pathologists : SNOMED Clinical Terms Guide. Version 5.0, 2002-2006