

概念マップによるオントロジー構築・洗練支援に関する基礎的考察

Fundamental Consideration for Supporting Ontology Construction/Refinement with a Conceptual Map

太田 衛*¹
Mamoru Ohta

古崎 晃司*¹
Kozaki Kouji

溝口 理一郎*¹
Richiro Mizoguchi

*¹ 大阪大学産業科学研究所

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

When ontology developers build domain ontology, it is necessary for them to work in partnership with its domain experts. This paper discusses how the users make a conceptual map which visualizes ontology according to their own perspectives in ontology development process. We also consider how the conceptual map makes easier for domain experts to understand contents of ontology and how the tool supports ontology construction and refinement with using conceptual maps.

1. はじめに

近年、専門領域の情報化が進み、専門知識が細分化される中で、様々な専門領域で知識の構造化へのニーズが顕在化しつつある。知識の構造化を行う手段のひとつとしてオントロジー工学が注目されており、様々な領域を対象としたオントロジーの構築が進められている。このような背景のもと、オントロジーの専門家のみならず、各ドメインの専門家が適切なオントロジーを構築できるよう支援するための方法論やツールの開発が求められている。オントロジー構築のプロセスは設計、構築、洗練の3つのフェーズに大きく分かれるが、オントロジーが対象世界を適切に表しているかを評価し、その結果に基づいて洗練を行う洗練フェーズは、オントロジーの質を保証する上で重要な位置を占める。

一般的にオントロジーの評価としては、形式的な評価と内容的な評価の2つの評価に大きく分けられる。形式的なアプローチに関しては、OWL では記述論理(DL)による推論処理で矛盾を検出するなどの機械的なアプローチをとることが一般的である。一方、内容的(意味的)な評価に関しては、構築したオントロジーに基づいたモデルを実際の対象世界の問題解決に適用して有効性を示す方法や、ドメインの専門家による内容評価を行う組織を形成するなど人的なアプローチを取る方法がある[Barry 07]。オントロジーの質を保証するには、ドメイン知識をもつ専門家の協力を得て内容評価を行うことが必須ではあるが、オントロジーの評価には、ドメインの専門知識もさることながら、オントロジーの表現内容を理解する能力が必要となってくる。特にオントロジーに対するリテラシの乏しいドメインの専門家にとって、オントロジーの定義内容を見て意味的な確認を行う作業は容易ではない。

そのような問題に対して、ドメインの専門家がオントロジーの記述内容を理解することを支援する為のアプローチとしては、オントロジーの可視化によって直感的に理解しやすくする手法の検討[Katifori 07]や、統制言語を用いて自然言語に近い形で記述させ、それを計算機処理で形式的なオントロジー記述に変換するアプローチがなされている[Dimitrova 08]。筆者らもまたオントロジーと領域の専門家に関心のある知識の間の隔たりを埋め、オントロジーの内容を効果的に俯瞰する仕組みを提案している[廣田 09]。その取り組みの中では専門家の視点に応じてオントロジーから知識を抽出し、内容を把握しやすい概念マップ

の形式で表示する“概念マップ生成ツール”を提案しており、構築されたオントロジーと専門家の知識とのギャップを埋めるには、専門家の視点に応じて概念間の関係性を見せることが有効であるという知見が得られている。本研究では、オントロジーの構築プロセスの中で洗練フェーズに着目し、利用者の目的や視点に応じてオントロジーを探索・可視化する概念マップを用いて、ドメインの専門家がオントロジーの意味的な内容を理解して、洗練する過程を支援する枠組みについて考察する。

2. 概念マップによるオントロジー構築・洗練支援

2.1 オントロジー構築プロセスにおける位置づけ

オントロジー構築のライフサイクルは、大きく設計、構築、洗練の3つのフェーズに分かれる。

- オントロジーの設計(初期フェーズ)
ドメインの専門家とオントロジー専門家が協力して、対象領域における主要概念(知識)を抽出、整理し、上位概念構造の検討など、オントロジーの基本設計を行う。その過程では、概念定義を行う上で何が本質的かといった本質属性や、概念の共通性・汎用性などを適切に捉える、オントロジー工学的な考察が重要となる。
- オントロジーの構築(中期フェーズ)
上位概念構造の設計がある程度固まってきたら、上位概念の定義に従って、各概念を増やしていき、ミドルレイヤー以降の概念構造を組織化する。
- オントロジーの洗練(後期フェーズ)
構築したオントロジーが対象とする世界を適切にモデリングできているかを、概念定義の内容を確認して評価し、問題点を修正することでオントロジーを洗練する。

本研究ではこの洗練(後期フェーズ)における構築・洗練支援を対象とする。

2.2 オントロジーの評価手法

オントロジーの評価には、形式的な評価と内容的な評価がある。この問題はソフトウェアの世界で言う verification(to make a program right)と validation(to make a right program)の問題と同様に考えることができる。オントロジーの評価をこの問題に当てはめると、以下ようになる。

- 形式的(文法的)に正しく書いているかどうかを評価する
 - 文法(表記法)に基づき正しく書いているかの検証
 - オントロジーとして各概念定義間の不整合性(矛盾)がないかの検証

- 内容的(意味的)に正しく書けているかどうかを評価する
 - ドメインの専門家による主観的な評価
 - オントロジー利用のユースケースに基づいた検証
 - オントロジーを実際の問題解決に適用した結果による評価
 - 過去の成功事例との比較(デザインパターンによる評価)

上記のうち、オントロジーの形式的な評価(verification)は推論機構などを用いた計算機処理で実現できるが、それでは無矛盾であることしか言えない。一方、その内容的な評価(validation)は現実世界のモデルを作ってみて、実例への適用や専門家の知識・経験を通してのみでしか判断できない。本研究では、上述のオントロジーの内容的な評価のうち、ドメインの専門家による主観的な評価を対象とする。

2.3 概念マップによるオントロジー構築・洗練支援

オントロジーを構築・洗練時に、ドメインの専門家がオントロジーの定義内容を適切に理解し、評価・洗練するには、オントロジーを専門家の視点でわかりやすく提示する必要がある。オントロジーの理解促進のための枠組みがあればドメインの専門家の理解や協力も得やすくなり、良質なオントロジー構築にも貢献すると思われる。更に概念間の関係性を現実世界のモデルに近い形で見せることができれば、実際のモデルと照らし合わせて確認出来るので内容評価しやすくなる。そこで本研究ではオントロジー構築・洗練の目的に応じてオントロジーを探索・可視化した概念マップによりオントロジーを洗練する過程を支援する枠組みについて提案する。

先行研究におけるオントロジーの俯瞰(利用)では、専門家の関心に応じて様々な視点からマップを生成する基本的な枠組みを提案してきたが、オントロジーの構築・洗練時には、その注目する視点が構築しているオントロジーや評価したい内容によって変わり、概念マップを生成する際に視点を絞り込む必要がある(図1)。このことはドメインの専門家は複雑な概念の関係性を、視点を切り替えることで対象を捉えていることと深く関連している。すなわち、概念マップによるオントロジーの構築・洗練の検討課題としては、評価したい内容に応じた概念マップ生成のために、どのような視点を選べば良いか、また抽出した概念間の関係性をどのように可視化すれば効果があるのか等の検討が必要とされる。

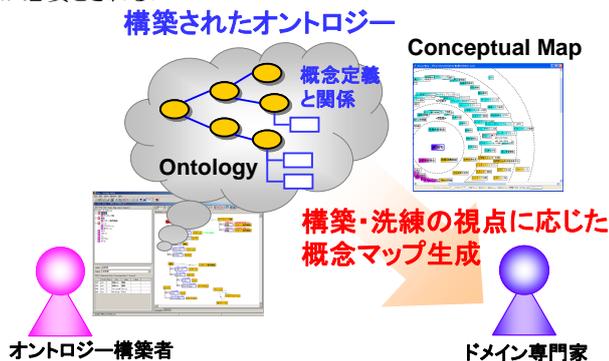


図1 概念マップによるオントロジー構築・洗練支援

2.4 概念マップのオントロジー構築・洗練の適用事例

前節で述べた概念マップによるオントロジー構築・洗練支援の検討課題に対して、サステナビリティ・オントロジーと臨床医学オントロジーを適用事例として、オントロジーの構築・洗練時に必要な概念マップを生成し、その有効性について考察した。

(1) サステナビリティ学オントロジー構築における適用

RISS(大阪大学サステナビリティ・サイエンス研究機構¹)における知の構造化研究グループでは、環境分野におけるサステナビリティ学に関する多様な領域の知識を、領域に依存しない形式で体系化したサステナビリティ学オントロジーの構築に取り組んでいる[Kumazawa 09]。その構築過程においては、構築したオントロジーから概念マップを生成し、対象とする知識が適切なマップとして生成できるかを確認することでオントロジーの構築・洗練を行ってきた。また同様な手法を、環境省「Hc-082 アジア太平洋地域を中心とする持続可能な発展のためのバイオ燃料利用戦略に関する研究」という別プロジェクトにおいてバイオ燃料分野に関連する概念を中心に既存のサステナビリティ学オントロジーを拡充する際にも適用した。その際には環境分野の専門家が作成した29の典型シナリオをもとにオントロジーを拡充し、それらのシナリオが概念マップとして再現出来ているかを確認することでオントロジーの洗練を行った。これにより既存のオントロジーとの整合も取りながら拡充したオントロジーの意味的な評価を行うことができ、この構築・洗練作業には有効であることが分かった。

これらの構築・洗練作業では、概念マップとして表したいシナリオが多様な内容であることから、概念マップ生成に用いる視点には特に制限を設けず、利用者が必要に応じて選択するという手法が求められた。

(2) 臨床医学オントロジー構築における適用

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」²では、臨床医学の複数領域にわたる電子データの統合管理や高度な情報解析を行う情報基盤技術として不可欠であるとして、東京大学・大江教授らの研究グループと大阪大学・溝口研究室との共同で、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]。臨床医学オントロジーの構築では、初期フェーズとして臨床医学の専門家と共同作業で臨床医学に必要とされる主要な概念が記述できるかどうかの検討を行い、具体的には疾患および疾患に関連する症状・所見、および人体構造知識を中心として、その基底となる上位概念構造の基礎的な考察を行っている[国府 08]。また中期フェーズとして、これら共通の上位概念構造をもとに、疾患および人体構造に関するオントロジーを専門家が構築している。

これらの構築・洗練作業では、ドメインの専門家がオントロジーの内容を確認する為に、共通の上位概念構造のうち注目した視点のみに限定して可視化した概念マップを生成するという手法が求められた。3章では臨床医学オントロジー構築における概念マップ生成の詳細について述べる。

3. 臨床医学オントロジー構築での概念マップ生成

3.1 着目した上位概念構造の関係性の可視化

2.4(2)項で述べたとおり、臨床医学オントロジー構築における概念マップ生成では、オントロジーの上位概念構造のある注目した視点での可視化が求められた。そこでオントロジー構築・洗練支援のための概念マップの生成方法として、臨床医学オントロジーの共通の上位概念構造を持つ概念群に対して、視点として着目した概念間の関係性だけを抽出して可視化する手法

¹ <http://www.riss.osaka-u.ac.jp/>

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

を採用した(図 2)。これにより、オントロジーの構築・洗練時にドメインの専門家が評価したい内容だけを可視化し確認しやすくなる。

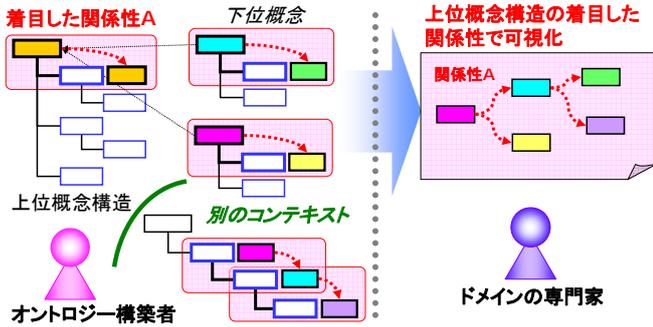


図 2 着目した上位概念構造の関係性に特化した可視化

3.2 人体構造の接続関係への適用

(1) 接続関係の定義

臨床医学オントロジーにおける人体構造の接続関係は、「ポート」という共通の上位概念構造を導入し、血管や神経といった流体や信号を通すような機能的な接続、骨と関節の接続といった力学的な接続、それ以外の位置関係を表すような空間的な接続¹など、様々な接続関係を同型のフレームで扱えるように工夫されている。これにより血管系や神経系といった機能的な接続関係の繋がりや骨どうしの繋がりを計算機で推論することが可能になっている。例えばある血管 A と血管 B の接続関係を、人体構造物とポートの概念を用いて表現した場合、接続ポートの接続先としてお互い接続ポートを相互参照する形で記述される(図 3)。接続の種類は、接続の記述に用いるポートの種類の違いで表している。さらに関節や血管の分岐点など、複雑な接続構造については関与する人体構造を仮想的な部品に分割するなどの手法を用いて、最小単位の接続に分解することにより、ポートを用いた同型の接続構造の組み合わせで捉えている。

現状の臨床医学オントロジーでは人体構造の接続関係について、心臓から器官までの血液の流れを対象とした循環器系統

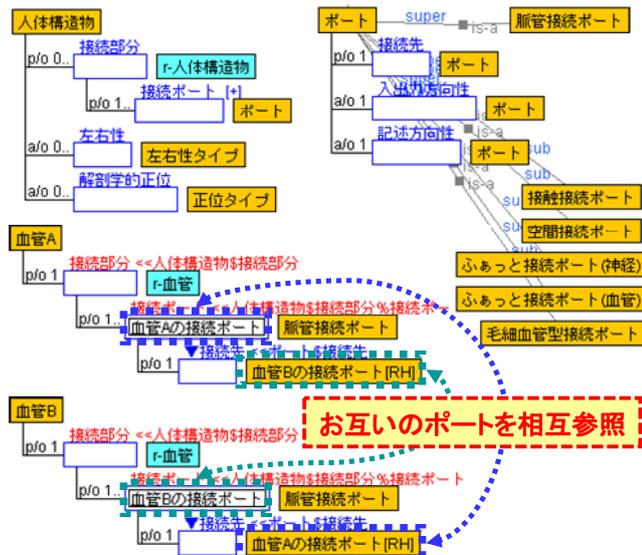


図 3 人体構造における接続関係の記述

¹ 位置関係や空間的に近接しているといった本来は接続とは言えない概念も、共通して枠組みで捉えるためにポートを用いた同形の概念構造で捉えている。

(動脈系、静脈系)、全身の主要な神経系統、骨格筋の接続関係のオントロジーが構築されている。各接続関係のオントロジーで定義されている概念数は、動脈系、静脈系を合わせた循環器系統で約 11,000、神経系統で約 7,500、またそれらの間の接続関係の数は、それぞれ約 8,600、約 3,200 となっている。

(2) 人体構造の接続オントロジーの構築・洗練支援

3.2(1)項で述べた人体構造の接続オントロジーにおいて、接続関係が内容的に正しく記述されているかどうかを確認するには、人体構造物の接続ポートの接続先に記述されている内容を順番に辿って確認すればよい。例えば、図 3 で示した血管 A と血管 B が接続されているかを確認するには、{血管 A → A の接続部分 → A の接続ポート → B の接続部分 → 血管 B} と辿ればよい。すなわち、循環器系統において心臓から送り出された血液が動脈の血管を通して全身の器官や組織に供給され、静脈の血管から心臓に戻ることが正しく表されていることを確認するには、このようなポートに沿って接続関係を辿る作業を接続の数だけ繰り返す必要がある。

しかし現実的な問題として、1 万を越える人体構造の接続関係を概念定義の内容を手動で辿って確認するのは事実上不可能に近い。よって、このような人体構造の接続関係が正しく記述されているかを確認し、オントロジーを洗練するには、必要な探索を自動的に行い、利用者(専門家)が把握しやすい形式で可視化することが求められる。しかもこのような支援は、共通性のある概念定義の場合には機械的な推論処理での対応が可能な範囲であり、その作業負担の軽減や利用者の内容理解に対する貢献度も高い。

そこで、このような必要性に応えるため、人体構造を記述したオントロジーから接続関係を表す「接続ポート」を辿るような探索を行い、その探索結果を人体構造の接続関係の確認用に可視化する専用ツールを試作した(図 4)。この可視化ツールの基本操作として起点となる人体構造物を選択し、そこから画面右下のリストに表示される接続先の人体構造物を選択することで、接続関係を辿って逐次表示していく。例えば、図 4 の画面例では、「心臓」を起点とする「大動脈」の接続関係を可視化しており、「心臓」から「脈管接続」を辿ると、「大動脈分岐」で分岐して「大腿動脈」まで「動脈」が繋がっていることが可視化によって理解できる。また「右外腸骨動脈」からは「ふあっと接続²」を介して「右大腰筋」といった筋組織へと血液が供給されていることもわかる。

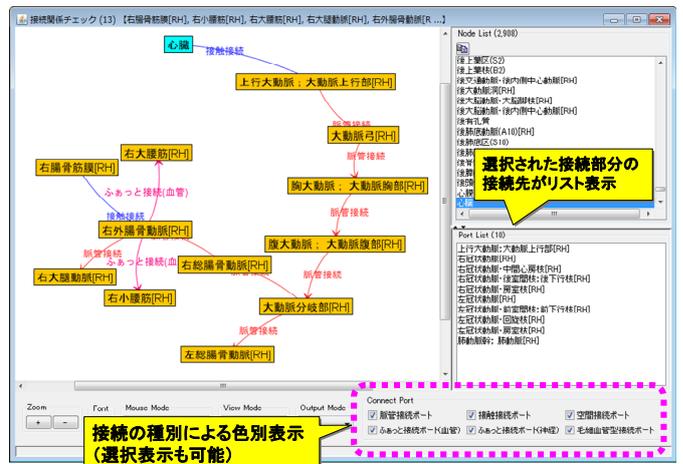


図 4 人体構造の接続関係の可視化ツールの画面例

² 血管から毛細血管を解して各組織へ血液が流れていることを表す接続

この可視化ツールでは、接続関係の種類はリンクの色を変えることで示すと共に、必要に応じて表示する接続関係の種類を切り替えることができる。また、指定した起点からすべての接続を再帰的に辿るといった、探索・可視化方法の選択ができる。

このように臨床医学オントロジーの人体構造物の接続関係の可視化を行うことによって、記述内容が直感的に理解しやすくなり、概念記述のエラーも見つけやすくなった。実際に専門家と共に「動脈」の接続関係を、可視化ツールを使って確認したところ、本来繋がっているべき血管の接続関係が記述できていないことが発見できた。

なお、可視化のための関係性の探索は、3.1節で述べた考え方に基づいて3.2(1)項で述べた人体構造における接続関係の共通の上位概念構造を用いて視点に基づき、関係の種類によらず統一した処理で実現されている。また、この接続関係の可視化ツールはJavaのクライアントアプリケーションとして開発し、概念間の関係ネットワークの可視化には、オープンソースのグラフ描画ツールJUNGを利用している[JUNG]が、オントロジー構築・洗練支援のためにより最適な可視化方法については検討を継続している。

3.3 疾患概念への適用

3.2節で試作した可視化ツールを臨床医学オントロジーにおける疾患概念の構築・洗練にも適用して検証した。

臨床医学オントロジーにおける疾患は、「その原因と途中経過をふくめた一連の状態変化と、それにより引き起こされている結果状態との総和」として捉えられ、疾患概念はその疾患を発症した患者には共通して見られる状態変化(連鎖)と定義される。また、疾患を構成する「異常状態」は患者に起こりうる状態として一般化されたものであり、「原因」と「結果」の二つの異常状態を属性として持つ概念として定義される(図5(a))。例えば、疾患「糖尿病」に共通して見られる異常状態「高血糖状態」の概念定義は図5(b)のようになる。疾患「糖尿病」に定義において、患者には共通して見られる状態変化は記述されているものの、その疾患を発症した患者に起こりうる異常状態の変化に関しては、汎用的な異常状態の「原因」及び「結果」属性を辿ることによって得られる。例えば「高血糖状態」の結果として「網膜血管障害状態」があり、その結果として「失明状態」があると「結果」属性を辿ることによって、その状態変化を得ることができる。このように疾患で定義された異常状態およびその状態変化を汎用化して集めることによって、患者に起こりうるすべての異常状態を得ることが出来る。

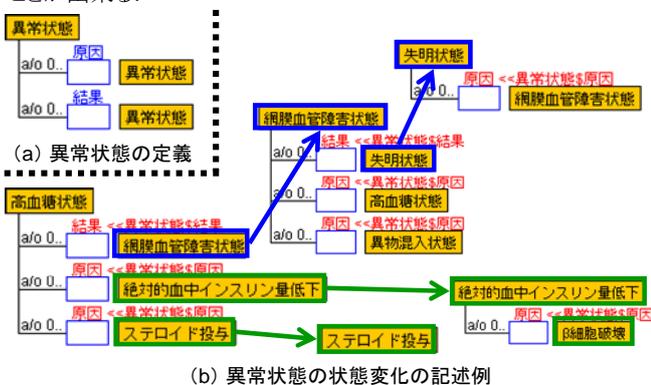


図5 疾患における異常状態の連鎖の記述例

現状の臨床医学オントロジーでは、疾患概念について12診療科の疾患が定義されており、その概念数は合計約6,000となっている。この疾患概念の洗練では、3.2(2)項と同様の仕組みで、疾患概念のオントロジーからすべての起こりうる異常状態の

状態変化を探索し可視化するツールを試作し、疾患概念の定義内容に加えて起こりうる異常状態の状態変化も確認出来るようにした。

4. まとめ

本研究では、オントロジーの構築・洗練において、ドメインの専門家に対して、その目的に応じてオントロジーを探索・可視化した概念マップを生成し、オントロジー構築・洗練を支援する枠組みについて提案し、実際のオントロジー構築での適用事例を用いて考察した。サステナビリティ・オントロジーでの構築・洗練では、概念マップを利用することで、既存のオントロジーとの意味的な整合も取りながら確認することができた。臨床医学オントロジーでの構築・洗練では、視点に応じた探索に加え、確認したい内容に合わせて可視化方法を変えることで有効性を示すことができた。今後の検討課題として、構築・洗練を目的としたマップ生成に関しては、オントロジーの上位概念の定義内容に強く依存しており、一般化・パターン化が困難であると考えられる。現実問題としての重要性がある為、まずはドメインに限らず成功事例を増やしていきながら検討を進めていくことが今後の発展の鍵となると考えている。そのなかで洗練タスクの概念化、上位概念構造による概念マップ生成、概念マップの視覚化方法といった各要素を切り分けて議論することが必要である。また、概念マップ上で発見した概念間の関係性の矛盾に関しても、何らかの編集支援が必要であり、今後の検討課題の一つである。

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19-21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

[Barry 07] Barry, S. et al.: The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration, *Nature Biotechnology* 25, 1251 – 1255, 2007

[Dimitrova 08] V. Dimitrova, R. Denaux, G. Hart, C. Dolbear, I. Holt and A. G. Cohn: Involving Domain Experts in Authoring OWL Ontologies, *Proc. of 7th International Semantic Web Conference, ISWC-2008, Karlsruhe, Germany, October 1-16, 2008*

[廣田 09] 廣田健, 古崎晃司, 斉藤修, 溝口理一郎: ドメイン知識俯瞰のためのオントロジー探索ツールの開発, 第23回人工知能学会全国大会, 2009

[JUNG] <http://jung.sourceforge.net/>

[Katifori 07] A. Katifori, C. Halatsis G. Lepouras, C. Vassilakis and E. Giannopoulou: Ontology Visualization Methods—A Survey, *ACM Computing Surveys (CSUR), Surv. 39, 4, Article 10, 2007*

[国府 08] 国府裕子, 周俊, 古崎晃司, 今井健, 大江和彦, 溝口理一郎: 臨床医療オントロジーの構築に関する基礎的な考察, 第22回人工知能学会全国大会, 2008

[Kumazawa 09] T.Kumazawa, O.Saito, K.Kozaki, T.Matsui, R. Mizoguchi: Toward Knowledge Structuring of Sustainability Science Based on Ontology Engineering, *Sustainability Science, Vol.4, No.1, 2009.*

[Mizoguchi 09] R. Mizoguchi, H. Kou, J. Zhou, K. Kozaki, K. Imai and K. Ohe: An Advanced Clinical Ontology, In *Proc. of International Conference on Biomedical Ontology (ICBO), pp.119-122, Buffalo, NY, June 24 – 26, 2009*