

多段階展開型オントロジー内概念検索システムの試作

A Prototype of Multi-Level-Expansion Search for Concepts in Ontology

北河 祐作
Yusaku Kitagawa

古崎 晃司
Kouji Kozaki

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 (ISIR), Osaka University

Recently, various ontologies and applications based-on them are developed. For ontology-based applications, it is very important to acquire appropriate concepts requested by the system developer. In this article, we propose "Multi-Level-Expansion Search" for properly acquisition of concepts from ontologies according to the developer's intentions.

1. はじめに

高度な意味処理を計算機上で実現するために、オントロジー工学の研究が進められており[溝口 05]、様々な領域においてオントロジーの構築や、構築されたオントロジーを用いた知識システムの開発も進められている。

オントロジーベースの知識システム開発において、オントロジーに記述されている内容を適切に利用するためには、システムが必要とする情報をオントロジーから正確に取り出すことが不可欠である。すなわち、オントロジーを対象に行う意味処理の中でも、オントロジー内に定義されている概念から、概念定義(意味)に基づく検索は重要な要素技術であることがわかる。オントロジー内で定義されている概念を検索する際には、オントロジーの概念的性質上、様々なアプローチが存在するが、現状の検索手法の多くは単に指定された検索条件を満たす概念をすべて返すといった最も基本的な手法に限られている。しかし、ユーザ(システム開発者)が概念を検索する意図には、より多くの観点が考えられる。例えば「骨の疾患」を検索する際には、「どのような骨にも起こりうる疾患」を知りたい場合や「肋骨や軟骨などの具体的な骨に起こりうる疾患」を知りたい場合があり、「骨の疾患」だけではどのような意図の検索であるか不明確である。さらに、検索対象のオントロジーは、特定の領域を対象に構築されたオントロジーだけではなく、どのようなオントロジーに対しても、汎用的に検索が可能でなければならない。

以上より、オントロジーの領域に依らず共通して利用可能であり、ユーザの検索目的に応じた検索手法が必要であることがわかる。そこで本研究では、ユーザの検索目的に応じて適切な概念を段階的に検索する「多段階展開型検索手法」を提案する。

2. オントロジー内概念検索

オントロジーを用いた意味処理において、オントロジーで定義された概念構造に基づき適切に概念や関係を取り出すことは、オントロジーや意味処理の種類に依らず共通する要素技術となる。これは、意味処理の対象となる内容を適切な検索条件として表し、オントロジー内からその条件を満たす概念(や関係)を検索することに相当する。本研究では、このような「オントロジー内概念検索」を対象として、高度なオントロジー利用に向けた適切な検索手法について考察する。

本章では、検索対象をオントロジー内で定義されている概念

連絡先: 北河祐作, 大阪大学産業科学研究所, 〒567-0047
大阪府茨木市美穂ヶ丘 8-1, Tel: 06-6879-8416, Fax:
06-6879-2123, E-mail: kitagawa@ei.sanken.osaka-u.ac.jp

とし、オントロジーの種類に依らず適用可能な、概念の定義内容に基づいた検索の基本要素について考察する。

2.1 基本的な検索要素

オントロジーにおいて、概念は他の概念との関係性により定義される。よってオントロジー内の概念を対象とした意味検索は、単なるラベル一致を条件とした文字列による検索ではなく、概念間の関係を用いた検索を指す。オントロジーの概念定義に用いられる関係には、概念間の一般-特殊関係を表す is-a 関係と、各概念の定義内容を表すのに用いられるその他の関係(part-of 関係, attribute-of 関係など)がある。オントロジーの定義内容を利用することで、概念又はスロットの検索が可能となる[北河 11]。まず、検索対象が(基本)概念である場合に、検索条件に用いられる要素は、

- c1) 概念名(ラベル)
- c2) is-a 階層(上位/下位概念の指定)
- c3)(スロットで表される)定義内容

であり、後者2つが意味処理の中心となる。

次に、検索対象がスロットである場合に、検索条件に用いられる要素は、

- s1) ロール概念
- s2) クラス制約
- s3) ロールホルダー
- s4) 個数制約
- s5) スロットの種類(p/o もしくは a/o)
- s6) スロットの定義内容

であり、以上が、オントロジー内の概念を検索する際の最小単位となる基本的な検索要素である。実際に検索する際は、これらの基本要素を組み合わせることで、より高度な検索が実現される。

2.2 基本的な要素の組み合わせ

オントロジー内概念の検索において、意味処理の対象とする概念を取り出す際に必要とされる最も基本的な検索条件は、「指定した概念定義(意味)を満たす概念の検索」である。これは前節で述べた検索条件のうち、c3)「指定した定義内容による検索」に相当し、定義内容は前節で示した s1)~s6)の条件を用いて指定される。その中でも s1)「ロール概念」と s2)「クラス制約」の組み合わせを指定したスロットの検索が頻繁に行われる。これは OWL では「指定したプロパティの種類と値域(range)を Restriction として持つ Class」の検索に相当する。

例えば疾患オントロジー[Mizoguchi 09]において、「骨の疾患」を検索したい場合は、「ロール概念が「異常対象」でクラス制

約が「骨」であるスロットを定義内容として持つ概念”を検索すればよい(図 1)。このように、概念定義から概念を検索する場合、オントロジー内の概念が厳密に定義されている高級なオントロジーを対象にすることで、様々な検索が可能となる。

本研究では、基本的な検索要素の様々な組み合わせに関して考察するが、本稿では前述の c2)「指定した定義内容による検索」において s1)「ロール概念」と s2)「クラス制約」の組み合わせを指定した検索を中心に議論する。

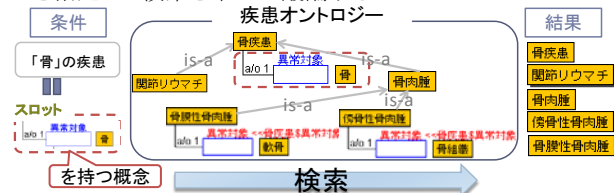


図 1 「骨の疾患」の検索例

2.3 問題点とアプローチ

オントロジーを用いた意味処理においては、オントロジー内の概念検索の結果が、ユーザ(システム開発者)の検索意図や目的を適切に反映していることが重要となる。例えば、図 1 の「骨の疾患」という条件で検索を行う際に考えられる意図としては、

1)「骨の疾患」に共通する性質(概念定義)を知るため、検索結果の中でも最も一般的な疾患の検索を目的とし、検索結果として「骨疾患」のみを求める場合

2)(後頭骨や脊椎骨や肋骨のような特定の種類の骨にのみ起こる疾患ではなく)すべての種類の骨に共通する疾患の検索を目的とし、検索結果として異常対象のクラス制約が骨(特定の種類の骨は除く)である「骨疾患」、「関節リウマチ」、「骨肉腫」を求める場合

3)後頭骨や脊椎骨や肋骨のようなあらゆる種類の骨を分類した部位における疾患の検索を目的とし、図 1 で示した全ての疾患を検索結果として求める場合

などがある。“ロール概念が「異常対象」でクラス制約が「骨」であるスロットを持つ概念”という「検索条件を満たす」という点だけを考えると、図 1 で示した全ての概念を検索結果として返すべきである。しかし、1)~3)のようなユーザの意図を考慮すると、検索結果として返すべき範囲が異なり、単に上述の条件を満たすというだけでは、ユーザの意図が不明確である。また検索結果の中にユーザが求める情報が含まれていないとしても、膨大な検索結果が返ってきた時などには、求める結果の有無がユーザにとって不鮮明である場合も考えられる。

そこで本研究では、ユーザの意図に応じた検索に対応するために、概念の is-a 階層をユーザの意図に応じて段階的に展開することによる多段階展開型検索手法を提案する。

3. 多段階展開型検索手法

オントロジー内概念検索において「指定した定義内容による検索」を行う際には、検索条件としてスロットで表わされる概念定義を指定する。この検索条件と、その条件から得られる結果のそれぞれに対して is-a 階層に沿った下位概念へ段階的に展開を行うことで、ユーザの意図に応じた検索を実現する。さらに、それぞれの展開を組み合わせることで、検索手法を複数の段階に細分化し、それぞれを検索目的の範囲に応じて区分する。

3.1 結果展開

検索条件を満たす結果として得られた概念の下位概念には、is-a 関係に沿って定義内容が継承されるため、それら全ての下

位概念も検索条件を満たす。このように、検索結果として得られた概念の is-a 階層に沿って下位概念を展開し、それらの下位概念も検索結果とすることで、結果として返す範囲を操作することを結果展開と呼ぶ。例えば、様々なスポーツに関して定義されている「スポーツオントロジー」において「ボールを用いるスポーツ」という条件で検索した場合、結果展開を適用することで、「球技」の下位概念である「野球」や「サッカー」なども結果とする。

さらに、スロットの継承・特殊化を考慮することで、結果展開を 3 レベルに区分することが可能となる。スロットの特殊化とは、下位概念において上位概念から継承された定義内容をより詳細なものに書き換えることである。図 2 の「大動脈疾患」の例では、ロール概念が「異常対象」でクラス制約が「大動脈」であるスロットが定義されている。その下位概念である「大動脈炎症候群」と「大動脈瘤」にはそのスロットがそのまま継承されており、さらに下位概念である「腸骨動脈瘤」、「腹部大動脈瘤」、「胸部大動脈瘤」では、継承されたスロットの「異常対象」のクラス制約が「大動脈」から「腸骨動脈」や「胸部大動脈」に特殊化されている。なお、スロットが特殊化されていない場合も、上位概念から継承された定義内容は満たしているため、検索結果としての条件は満たしている。このスロットの継承・特殊化を考慮した結果展開のレベル分けを以下に示す。

結果展開:無し

- ・結果展開を用いない
- ・検索条件を満たす最上位概念のみ取得

結果展開:レベル 1

- ・検索条件となるスロットをそのまま継承している概念まで展開
- ・該当スロットが特殊化されている概念は展開しない

結果展開:レベル 2

- ・検索条件となるスロットが特殊化されている概念も展開
- ・レベル 1 で得られる概念の下位概念全てを展開

図 2 のオントロジーで、「大動脈の疾患(「異常対象」のクラス制約が「大動脈」である疾患)」という条件で検索する際に、結果展開無しでは「異常対象」のクラス制約が「大動脈」であるスロットが定義されている「大動脈疾患」のみを結果とする。結果展開レベル 1 を適用すると、検索条件と一致するスロットを継承している「大動脈炎症候群」と「大動脈瘤」も結果とし、さらに結果展開レベル 2 を適用すると、そのスロットが特殊化されている「腸骨動脈瘤」、「腹部大動脈瘤」、「胸部大動脈瘤」も結果とする。

また、スロットの特殊化は、クラス制約の特殊化とロール概念の特殊化に分けられる¹。例えば、図 2 のようにクラス制約のみが特殊化される場合と、図 3 のようにクラス制約は「人間」のまま継承し、ロール概念が「構成員 role」から「部員 role」や「社員 role」に特殊化される場合がある。図 2 や図 3 のように、クラス制約とロール概念のいずれかが特殊化されている場合、クラス制約に着目した結果展開レベル 1 とロール概念に着目した結果展開レベル 1 に差異が生じる。例えば、図 3 において“「構成員 role」のクラス制約が「人間」である概念”を検索する場合、クラス

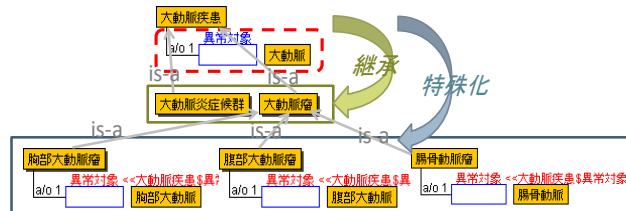


図 2 スロットの継承・特殊化(クラス制約)

¹ クラス制約とロール概念以外に、個数制約やロールホルダーなどについても同様の議論が可能であると考えられる。

制約に着目した結果展開レベル 1 では、「人の集まり」、「組織」、「部署」、「会社」が検索結果として得られる。一方でロール概念に着目した結果展開レベル 1 では、「人の集まり」、「組織」が検索結果として得られる。よって、クラス制約とロール概念いずれかに着目した結果展開の適用が可能であることがわかる。また、クラス制約とロール概念を複合条件とした結果展開レベル 1 の場合、検索結果はそれぞれの結果展開レベル 1 の検索結果の論理積をとることが妥当であると考えられる。そうすることで、「検索条件となるスロットをそのまま継承している概念まで展開」という結果展開レベル 1 の定義を満たすことができる。

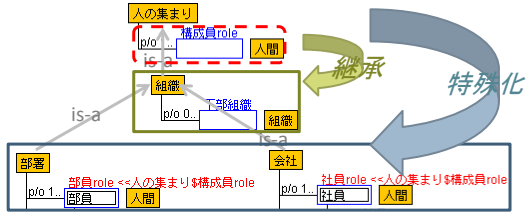


図 3 スロットの継承・特殊化(ロール概念)

3.2 条件展開

検索条件として「クラス制約で参照している概念」を指定したとき、「その下位概念をクラス制約で参照している」場合も、オントロジー的に条件を満たしている。よって、「指定したロール概念とクラス制約の組みを定義内容にもつ概念」を検索する際には、検索条件で指定したクラス制約で参照している概念の is-a 階層を用いて、その下位概念をクラス制約で参照して検索条件で指定したロール概念を定義内容に持つ概念も検索結果に含まれる。このように検索条件で参照している概念の is-a 階層を展開し、検索条件を満たす概念を検索する手法を条件展開と呼ぶ。例えば、「ボールを用いるスポーツ」という条件で検索するとき、ロール概念「使用道具」のクラス制約が「ボール」であるスロットを定義内容に持つ概念を検索する。その際に、条件展開を用いずに検索すると、「使用道具」のクラス制約が「ボール」と定義されている概念である「球技」のみを検索結果とする。一方、条件展開を適用すると、クラス制約で参照している「ボール」の is-a 階層を用いることで、「ボール」の下位概念である「野球ボール」や「サッカーボール」を「使用道具」としてクラス制約で持つ「野球」や「サッカー」なども検索結果として得られる。このように、条件展開を用いることで、一般的な条件での検索と詳細な条件での検索に段階付けることが可能となる。

また、ロール概念に関しても is-a 階層が存在するため、クラス制約で参照している概念と同様に、条件展開が可能である。ロール概念階層による条件展開により、検索条件で指定したスロットのロール概念が特殊化されている概念を得ることができる。

なお、条件展開では、検索結果となる概念の is-a 階層、すなわち、上位概念から継承された定義内容は考慮しない。よって、上位から継承されたスロットが検索条件を満たしていたとしても検索結果には含まない。

3.3 その他の is-a 関係の結果展開・条件展開

(1) p-is-a 関係の条件展開

part-of関係に沿った定義内容の特殊化が考えられるとき、その定義内容が参照しているのは「参照している概念そのもの」ではなく、参照している概念の「部分性のみに着目して抽象化した概念」であると考えられる。法造¹ではこのような「部分性のみに着目して抽象化した概念」を「p-オペレータ」を付与することに

よって表現する[古崎 10]。それにより、part-of関係をis-a関係に読み替えて得られた、概念間の全体-部分関係を表すp-is-a関係が構築される。このようにp-is-a関係は、クラス制約にp-オペレータを用いた際にシステムが動的に生成するis-a関係であるため、多段階展開型検索手法において、p-is-a階層は条件展開のみに用いられる。

(2) IS-A 関係の結果展開・条件展開

IS-A 関係は「identity の継承を伴わない属性継承のみを伴う弱い is-a 関係」を表し、法造においては、a) 多重継承を表すためにオントロジー構築者が明示的に記述した IS-A 関係と、b) スロットのロール概念定義に伴い導出されるプレイヤーからロールホルダーへの属性継承を表す IS-A 関係、の 2 種類に大きく分けられる。[太田 11]

a) に関しては、属性継承を表す IS-A 関係であるため、is-a 関係と同様に、結果展開と条件展開に適用できる。

b) は<ロールホルダー IS-A プレイヤー>と表現される。これに伴う属性継承は、そのロール概念が依存するコンテキストの下でのみ成り立つ。よって、is-a 関係やその他の IS-A 関係のような属性継承だけではなく、概念(ロールホルダー)のコンテキストが大きく影響する。そのため、結果展開を適用する際には、コンテキストに依存している概念を検索結果とすると、元の検索条件の目的から逸脱することが考えられる。例えば、得られた検索結果である「動物」から結果展開を適用すると、<教師 IS-A 人間 is-a 動物>という is-a(IS-A)階層に沿って、「教師」も検索結果となる。「教師」は「学校」というコンテキストに依存しているため、検索結果としては不適切であると考えられる。従って、b) による結果展開は、その他の結果展開と峻別する必要がある。一方で、条件展開に関しては、峻別する必要がない。なぜなら、検索条件の目的は用途(ロール)を考慮しないためである。例えば、図 4 のように、「ラジエータ」の定義内容より、<リザーバータンク RH IS-A 貯水タンク>であることがわかる。そして、「構成要素」として「貯水タンク」を持つ概念を検索する際に、条件展開を適用すると、「貯水タンク」がどのような用途(ロール)で用いられていることに関係なく、「リザーバータンク RH」を含む全ての下位概念を検索条件に適用する必要があると考えられる。

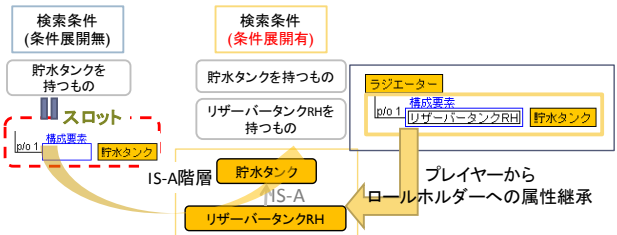


図 4 IS-A 関係による条件展開

3.4 条件展開と結果展開の組み合わせ

(1) 結果展開/条件展開で得られる検索結果の違い

結果展開および条件展開を適用した際に得られる検索結果の違いについて、それぞれの展開で利用する is-a 階層構造に注目して考察する。

結果展開のみを適用した場合は、元の検索条件を満たす概念のみを検索結果とする。一方、条件展開では、元の検索条件で用いられる概念の下位概念も、新たな検索条件として用いられる。そのため、元の検索条件を満たす概念の下位概念以外であっても、新しい検索条件を満たしていれば条件展開を適用した場合の検索結果となり得る(図 5-[a])。すなわち、条件展開により得られた新たな検索条件を満たす概念が、元の検索条件

¹ <http://www.hozo.jp/>

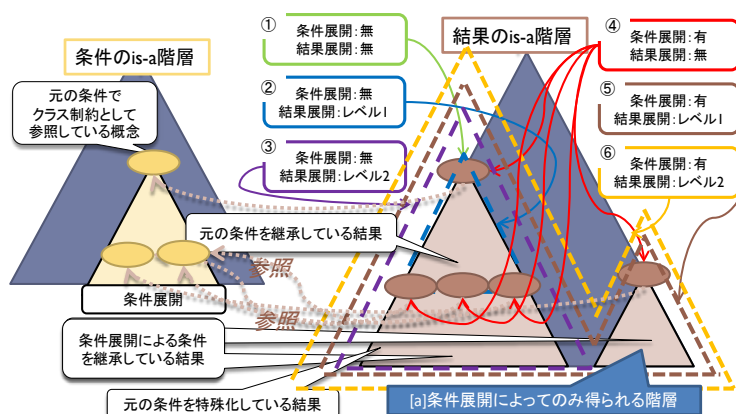


図5 展開パターン毎の検索結果の適用範囲

を満たす概念の下位概念以外の階層に現れる場合、結果展開では得られない検索結果が得られることがわかる。このような結果展開と条件展開の差異は、オントロジーを構築する際に、検索条件で参照される概念の is-a 階層と、検索結果となる概念の is-a 階層の分類視点が異なる場合、すなわち、それらの is-a 階層が相似形になっていない場合に生じる。

(2) 組み合わせによる展開パターンと検索結果の範囲

本項では、条件展開と結果展開の組み合わせによる展開パターンについて考察を行う。条件展開は適用する/しない場合の2種類、結果展開は適用しない/レベル1/レベル2の3種類存在するので、展開パターンの組み合わせは6種類存在することがわかる。それぞれの展開パターンを適用した際に検索結果として得られる範囲については図5のようになる。

まず、条件展開を用いない①～③の展開パターンは、3.1節で述べた3種類の結果展開と同様であるので、④～⑥の展開パターンについてのみ以下に示す。

④「条件展開:有り, 結果展開:無し」

④では、元の検索条件により得られる概念および条件展開により得られる概念を検索結果とする。結果展開は、展開を適用せずに条件を満たす概念の最上位概念のみを検索結果とする結果展開無しであるので、条件展開によって得られた各階層のルートである概念が検索結果となる。

⑤「条件展開:有り, 結果展開:レベル1」

⑤では、④の条件展開の適用で得られた概念に加え、それらの検索結果に結果展開をレベル1(条件となるスロットをそのまま継承している概念まで展開する)で適用する。よって、⑤では範囲外であった検索条件となるスロットをそのまま継承している下位概念も検索結果となる。

⑥「条件展開:有り, 結果展開:レベル2」

⑥では、④の条件展開の適用で得られた概念に加え、それらの検索結果に結果展開をレベル2(下位概念全てを展開する)で適用する。よって、条件展開によって得られた各階層のルートである概念の下位概念を全て検索結果とする。

これら6種類の展開パターンで得られる検索結果の範囲を比較すると、⑤と⑥は常に同一となり、条件展開により得られる検索結果が、元の検索条件により得られる概念の下位概念から現れる場合においては、③、⑤、⑥で得られる検索結果の範囲も同一となる。つまり、得られる検索結果は最も多い場合で、①、②、③、④、⑤-⑥の5パターンとなることがわかる。

3.5 展開パターンの段階区分の考察

次に5種類の展開パターンを検索目的の範囲に応じた段階区分の考察を行う。

①「条件展開:無し, 結果展開:無し」

①では、結果展開と条件展開共に適用しないため、元の検索条件であるスロットが定義されている概念のみを検索結果とする。そのため、①の検索は、「検索条件を満たす最も一般的な概念を検索」する際に用いられると考えられる。

②「条件展開:無し, 結果展開:レベル1」

②では、元の検索条件であるスロットが特殊化されずに継承している概念を展開し、検索結果とする。②の検索は、「検索条件を満たす最も一般的な概念を検索条件として注目している性質(概念定義)を特殊化せずに継承している概念を検索」する際に用いられると考えられる。

③「条件展開:有り, 結果展開:レベル2」

③では、②から元の検索条件であるスロットを特殊化している概念も展開し、検索結果とする。③の検索は、「条件の is-a 階層の分類の観点を満たす is-a 階層の概念を検索」する際に用いられると考えられる。

④「条件展開:有り, 結果展開:無し」

④では、元の検索条件であるスロットの特殊化が定義されている概念を検索結果とする。④の検索は、「検索条件として注目している性質(概念定義)が、詳細化される最上位の概念を検索」する際に用いられると考えられる。

⑤-⑥「条件展開:有り, 結果展開:レベル1/レベル2」

⑤-⑥では、従来の検索と同様に、オントロジー的に検索条件を満たす全ての概念を検索結果とする。⑤-⑥では、「検索条件を満たす全ての概念を検索」する際に用いられると考えられる。

4. まとめ

本研究では、ユーザの検索目的に応じて適切に概念を段階的に検索する「多段階展開型検索手法」を提案した。これにより、オントロジー内の概念を検索する際に、詳細に検索条件を設定するだけでなく、ユーザの意図に応じた5種類の展開パターンの提供が可能になると考えられる。また現在、提案手法を反映した検索システムの試作を進めている。

今後の課題として、検索結果を適切に提示するための検索結果提示方法について考察する必要がある。さらに、提案手法の展開パターンに対する検索意図の検証を行う。

謝辞

本研究の一部は科学研究費補助金若手研究(A) 20680009の助成による。

参考文献

- [Mizoguchi 09] R.Mizoguchi, H.Kou, J.Zhou, K.Kozaki, K.Imai and K.Ohe: An Advanced Clinical Ontology, In Proc. of ICBO, pp.119-122, June 24-26, 2009.
- [太田 11] 太田衛, 古崎晃司, 溝口理一郎:実践的なオントロジー開発に向けたオントロジー構築・利用環境「法造」の拡張—理論編—, 人工知能学会論文誌, Vol.26 No.2, pp.403-418, 2011.
- [北河 11] 北河祐作, 古崎晃司, 溝口理一郎:オントロジー高度利用のための多段階展開型検索手法の提案, 人工知能学会研究会資料, SIG-SWO-A1103-07, 2011.
- [古崎 10] 古崎晃司, 国府裕子, 今井健, 大江和彦, 溝口理一郎: 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—, 人工知能学会第24回全国大会, 1B5-1, 2010.
- [溝口 05] 溝口理一郎:オントロジー工学, オーム社, 2005.