

ユビキタス環境における概念階層と行動履歴を用いた意味層構築支援 Support of Semantic Layer Formulation by using Concept Hierarchies and Action History in Ubiquitous Community

三玉 政喜† 富井 尚志‡
Masaki Mitama Takashi Tomii

1 はじめに

近年、情報技術の発達により安価で高速な計算機やネットワーク、センサが普及し“いつでも・どこでも”計算機の恩恵を受けられる「ユビキタス環境」が現実味を帯び始めている。

また一方で概念を明示的にデータ化して、特定の分野の利用者間での意味統一を図るため、オントロジが導入されている。例えば Web 上のデータをマシンリーダブルにし、共有・検索をすることを目的とした Semantic Web がある。

それらをふまえて我々は、コミュニティごとに概念を明示化し、空間データに対応付けて管理・共有する、データベースを実装基盤とした概念共有環境 CONSENT を提案してきた[1]。この概念共有環境 CONSENT では、センサで取得した実世界情報をコミュニティ空間での共通知識、つまり「意味」や「意図」のような概念をデータベースに蓄積することでその共有・検索を可能とする。さらにユビキタス環境と VR を用いて、空間状態の取得や視覚的情報支援を行うことができる。

またこのようなユビキタス環境では無線 IC タグやさまざまなセンサを用いることで、実世界の「モノ」に関する移動の履歴の管理を実現できる。特に無線 IC タグを用いてデータ獲得した場合には、タグとリーダーの関係、すなわち「何を」「どこで」「どうする」を一意的 ID として取得可能である。一方でコミュニティ参加者（以下、利用者）共通の意味や操作の意図を表わす概念は、一般的にコミュニティごとに異なる。また人間の概念という性質上、計算機による自動構築も困難であるため参加者共有の概念を編集する人（以下、オントロジ編集者）が、手動で構築せざるを得ない。

ここで「何を」「どこで」は空間に存在する「モノ」なので一般によく用いられる物体に関するオントロジを利用することが可能である。一方「どうする」については、コミュニティごとに意図は異なるため、オントロジ編集者が手動で編集する必要がある。しかし、「何を」「どこで」「どうする」の組み合わせの候補数は膨大になり、その中から空間に存在する概念を選択するには困難を伴う。

そこで本論文では、無線 IC タグによって取得された利用者の操作履歴を利用して、オントロジ編集者がモノと操作との関係を示すオントロジを簡単に編集できるような手法を提案する。本手法では、ユーザの行動をトリガとして、行動履歴や概念階層を用いて、オントロジの候補となる要素を推奨順位付でオントロジ編集者に提示する。この推奨の手法として、本論文では以下の2つの手法を提案する。まず1つ目は、推奨しようとする関係オントロジと概念的

に近いと考えられる関係オントロジを参考に推奨する。2つ目は、関係オントロジのインスタンスである行動履歴から推奨する。この提案手法と protégé[2]等のオントロジエディタを併用することで、オントロジの構築コストの削減が期待できる。また行動履歴を用いて関係オントロジを推奨するので、空間の概念を反映した関係オントロジの構築が可能となる。

2 背景

2.1 オントロジ

オントロジとは本来は哲学用語で「存在論」を表すが、コンピュータ科学の分野では「概念化の明示的な規約」を表すものと定義されている[3]。

オントロジに関しては近年さまざまな分野で盛んに研究が行われている。World Wide Web では、情報にマシンリーダブルなセマンティクスを与えるための Semantic Web の開発が進められている。Semantic Web では情報をより詳細に記述するためにオントロジを用いており、オントロジを記述するための言語である OWL(Web Ontology Language)が2004年2月に W3C 勧告として公開された[4]。

しかしオントロジの構築には様々な問題があり、一般的にははまだあまり普及していないのが現実である。高梨ら[5]は、その理由としてオントロジの有効性の認知不足とオントロジの構築やメンテナンスのコストを挙げ、そのコストの削減のためユーザの操作を利用して構築のコストを削減する研究を行っている。

2.2 概念共有環境 CONSENT

ユビキタス環境では、人や物体の移動をセンサ等で検知することにより、空間情報を取得することが可能となる。オフィス、工場、学校や病院などのさまざまなモノや人が存在し、複数の参加者がそこで共通の意識を暗に共有している空間（以下、コミュニティ空間）では特にユビキタス環境による参加者支援が有効に活用されると考えられる。この「コミュニティ空間」をユビキタス環境による支援のために計算機上で再現しようと考えた場合、バーチャルリアリティによる仮想環境を用いる方法がある[6]。しかし実空間の物体に関する情報として位置情報や形状データの取得は可能であるが、実際にはそれらに加えて参加者間には様々な概念が暗黙のうちに存在している。

そこで我々は、コミュニティごとに概念を明示化し、空間データに対応付けて管理・共有する、データベースを実装基盤とした「概念共有環境 CONSENT(CONCEPT Sharing EnvironmeNT)」(以下、CONSENT)というユビキタス環境を提案してきた。この CONSENT ではコミュニティ空間での共通知識、つまり「意味」や「意図」のような概念を前述のオントロジとしてデータベースに蓄積することでその共有・検索を可能とする。さらに RFID などの無線 IC タ

† 横浜国立大学大学院環境情報学府
情報メディア環境学専攻
‡ 横浜国立大学大学院環境情報研究院

グや加速度センサなどを用いて、空間での物体の移動や人間の行動などの空間情報も取得することができ、VRによる支援も可能である。これにより「コミュニティ空間」の物体の位置や意味情報の検索、さらに人間の行動履歴の蓄積や空間内の概念の共有などが可能になる。

このデータベースのスキーマは空間に存在する意味を形状データと分離するため三層構造として構成している。意味を格納する「意味層」、実空間から得られるデータを格納する「マルチメディアデータ層」、その二つの層を結びつけマルチメディアデータが示す概念が空間に存在していることを示す「存在エンティティ層 (EE 層)」である。これらを独立して管理することで、概念の構造化を可能にし、概念を蓄積する意味層において一般に複雑な構造を持つオントロジの表現を実現する。

本研究で対象とする「意味層」は対象世界をモデル化するとき、オントロジを用いて空間の意味を明示的、体系的に抽出し蓄積する層である。CONSENT の意味層には物体に関する「物体オントロジ」、行動に関する「行動オントロジ」、そして2つの物体間の関係を記述する「関係オントロジ」がある。この3つのオントロジは現実空間の概念をモデル化した結果であり、この3つのオントロジによって空間の概念は自然に記述できる。また意味層のオントロジを記述できるのは、コミュニティ空間の概念を熟知しているオントロジ編集者のみとする。

3 行動履歴と概念階層を用いた意味層構築支援

この章では、概念共有環境 CONSENT における関係オントロジの構築コストを削減するためのオントロジ構築支援手法を提案する。

3.1 関係オントロジの構築コスト

物体オントロジおよび行動オントロジは、物体や行動の概念を記述したものである。CONSENT では名前やその説明、概念間の階層関係(is-a や part-of)を記述する手間が構築コストとなる。これらの物体や行動に関するオントロジは、外部のオントロジをインポートすることにより構築コストの削減が可能である。

ここで、実世界での利用者の行動を RFID によって獲得する場合には、IC タグの ID とそれを検知したリーダの組が取得される。この動作が仮に「何を(what)」「どこで(where)」「どうする(how)」を示しているとする。ここで what と where については物体に関するオントロジの要素

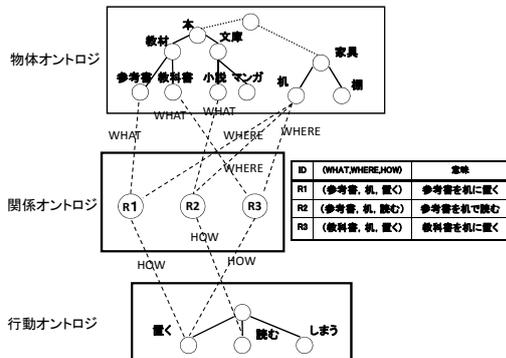


図1 意味層のオントロジの例

(WHAT,WHERE)を、how については行動に関するオントロジの要素(HOW)をそれぞれ対応付ければよい¹。

このとき明な組(WHAT,WHERE,HOW)は、このコミュニティ内で意味を持つ動作を表わすこととし、この組を関係オントロジと呼ぶ。関係オントロジの例を図1に示す。

また、関係オントロジを構築するコストは2つの段階に分割される。1つ目は関係がある2つの物体を選択する際のコストである。2つ目はその選んだ組の間にどのような行動ができるかを選択するコストである。物体オントロジの要素が n 個あったとき、2つの物体の組み合わせは n²通りあり、さらに行動オントロジ要素が m 個あると関係オントロジの候補は n²×m 通りある。その中から空間に存在している概念を人間が1つ1つ選ぶのには大きな手間がかかるといえる。本論文では、この関係オントロジの構築コストの両方の段階でそれぞれ候補を絞り込んで推奨することでコストを削減する手法を提案する。

なお、空間内での利用者の行動 R は、(WHAT,WHERE,HOW)の組で表現できないものも数多く存在する。しかし、この(WHAT,WHERE,HOW)の組は RFID を用いれば容易にセンス可能であることと、ならびに「モノ」の操作管理の面で有用であることから、本論文では特にこの組で表されるオントロジの編集を支援することを目指とする。またこの論文ではユーザによって行われ、センサ等で取得されたこの(who,what,where,how)の組を「行動」と定義し、行動オントロジとは別の意味で用いる。

3.2 概念階層と行動履歴

この節では推奨に利用する概念階層と行動履歴について詳しく述べる。

概念階層

概念階層とは一般的に分類階層、辞書が持つ語彙概念の階層などの階層知識を表し、クラスライブラリやオントロジといったものもこの概念階層を利用している。この概念階層は主に木構造で表される。

本論文では物体オントロジの持つ階層を概念階層と定義し、それを用いて推奨を行う。また図2のように、あるノードを基準として、同じ親を持つ兄弟ノードを Level1、親の親が同じ従兄弟ノードを Level2、といったように定義する。このレベルの値が小さいほど、基準ノードと概念的に近いと定義し、このレベルによって推奨の際にどの程度重要視して参考にするかを決定する。

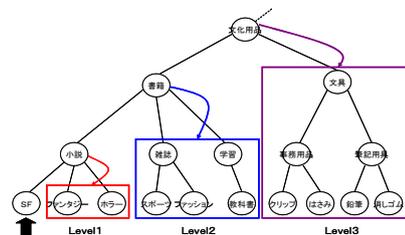


図2 概念階層

¹ センサ等で獲得された要素を小文字で、それに対応付けられるオントロジの要素を大文字で表現する。

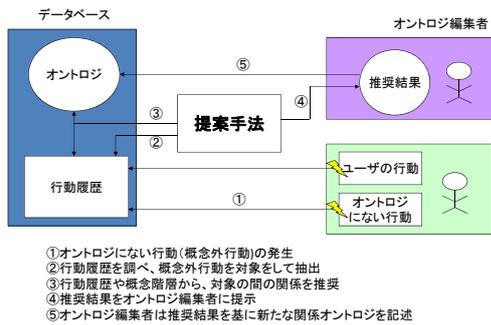


図3 フレームワーク

行動履歴

CONSENT では、RF-ID などのセンサや仮想空間のブラウザ上からの操作で空間での行動を取得し蓄積している。具体的にはAさん(who)が本 B(what)を机 C(where)に置いている(how)ということが蓄積されている。すべての行動を履歴とし蓄積したものをここでは行動履歴と呼ぶ。この行動履歴を用いて様々な支援が可能となる。しかしこれらの蓄積される行動は、関係オントロジに記述されていなければならない、それ以外の行動(以下、概念外行動)が行われたとき、行動履歴の代わりに関係対象が蓄積される。

3.3 関係オントロジの構築支援手法

オントロジの構築コストを削減するフレームワークは図3のようになる。まずユーザの行動をデータベースに蓄積していく。その中で概念外行動が生じ、その情報も前述のように行動履歴に蓄積される(①)。提案手法では一定期間ごとに行動履歴をチェックし、概念外行動を発見し、それを関係対象として抽出する(②)。すると提案手法では、概念階層や行動履歴を用いて関係対象の間で行える行動を推奨する(③)。その推奨結果をオントロジ編集者に提示し(④)、オントロジ編集者は推奨結果を基に新たなオントロジを記述する(⑤)。

このシステムを用いることで、表1のようなコスト削減が期待できる。さらにこのシステムでは、オントロジという「概念」の記述を行うので、最後のデータベースへの登録はオントロジ編集者が確認して行うことで、人間の意志を反映できる。

3.4 オントロジ推奨手法

本節では前述の提案システムのフレームワークの③にあたる、対象の間の関係を推奨する手法を2種類提案し、その手法について詳しく説明する。

RBO法

前述したように物体オントロジは概念階層で、木構造と仮定している。概念に基づいた木構造なので、ある概念ノードを基準とした兄弟ノードや従兄弟ノードは概念的に類

表1 関係オントロジ構築コスト

	支援無し	支援あり
対象(WHAT,WHEREの組)の選択	すべての物体オントロジから対象を選択	履歴から自動抽出
対象間の関係の選択	すべての行動オントロジから選択	推奨結果を基に選択

似していると考えられるのでそれらのノードに記述されている関係オントロジを用いて、関係対象の間の行動を推奨することができる。このように記述済みの関係オントロジを用いた推奨法をRBO(Recommended By Ontology)法と定義する。

いま関係対象(WHAT,WHERE)においてWHATをx, WHEREをyとするとき、xとyの関係オントロジの行動オントロジの集合をC(x,y)とする。さらにxを基準としたレベルをL_x, レベルがLのノードをN_L, レベルごとの重みパラメータをkとする。

$$C(x, y) = \{A_1, A_2, \dots, A_m\} \quad (1)$$

$$N_{L_x} = \{O_1, O_2, \dots, O_n\} \quad (2)$$

O_iを物体オントロジの要素, A_iを行動オントロジの要素とする。次にN_{L_x}の要素O_iのうち、あるyについて関係オントロジ(O_i,y,HOW)であるようなO_iの個数をoとする。ただしHOWは任意の行動オントロジの要素である。またそのうちHOWがある行動オントロジAであるようなO_iの数をaとする。例えばレベルLでのAに関する推奨度p_{A-RBO(L)}は

$$p_{A_RBO}(L) = \frac{a}{o} \times k^{(L-1)} + p_{A_RBO}(L-1) \quad (3)$$

ここで閾値のパラメータをtと指定すると、すべてのレベルの推奨度を加算した時の閾値T_Lは

$$T_L = k^{(L-1)} \times t + T_{L-1} \quad (4)$$

よって推奨度pがT_Lより大きい行動オントロジを、推奨度が高い順にオントロジ管理者に推奨する。この手法を用いることで、記述済みの関係オントロジを用いて新たな関係オントロジを記述でき、空間の概念にそった関係オントロジの記述が可能になる。

RBH法

CONSENTでは実世界の1つ1つの行動は空間センサにより取得されたEE(Existing Entity)としてDBに蓄積される。この履歴に空間の概念が含まれていることが多い。よって類似ノードのインスタンスに対して行われた操作や行動の履歴を用いたオントロジ推奨が可能となる。このように蓄積された履歴を用いた推奨法をRBH(Recommended By History)法と定義する。RBO法との違いとしてはRBH法ではインスタンスレベル、つまり実際に行われた行動のうち正しいデータを基に推奨を行うことである。

いま関係対象(WHAT,WHERE)においてWHATをx, WHEREをyとする。前節と同様にxを基準としたレベルをL_x, レベルがLのノードN_Lを式(2)のように定義し、物体オントロジO, 行動オントロジAを定義する。またここではレベルごとの重みはκ統計量を用いてκ_Lとする。いま履歴のなかでN_{L_x}のインスタンスi(N_{L_x})とyのインスタンスi(y)の間の行動履歴(i(N_{L_x}), i(y),how)であるような履歴の件数をhとする。howは任意の行動EEである。その中でhowのオントロジがAである件数をa'とする。その行動AにおけるレベルLでの推奨度p_{A-RBH(L)}は

$$p_{A_RBH}(L) = \frac{a'}{h} \times \kappa_L + p_{A_RBH}(L-1) \quad (5)$$

この手法を用いることで、蓄積された履歴を用いて新たな関係オントロジを記述できる。履歴はユーザが何らかの意図をもって実際に行った操作や行動の蓄積なので、それらはユーザの概念の集合であり、空間の概念が如実に現れていると考えられる。よって空間の概念を反映したオントロジの構築が可能となる。

κ 統計量

本論文では、RBH 法の場合にそれぞれのノードがインスタンスを持っていることを利用し、中間ノード同士の類似度に κ 統計量[7]を用いている。 κ 統計量では2つのノード間のインスタンスを表2のように区分している。i11 と i22 の数が多く、i12 と i21 の数が少なければ類似度は高くなる。

4 OntRecR の実装と評価

4.1 OntRecR の実装

本章では第三章で述べた提案手法に基づき、関係オントロジを推奨するツールである OntRecR を概念共有環境 CONSENT に実装した。OntRecR では主に3つの処理を行い、関係オントロジの構築支援を行なう。1つ目は履歴から関係対象を自動的に取得しオントロジ編集者に提示する。2つ目は選ばれた関係対象の間で行える行動を推奨する。そして3つ目は推奨された関係オントロジをオントロジ管理者の確認や変更を受けて DB に関係オントロジを登録する。履歴からの対象の抽出はオントロジ編集者が決めたスパンで履歴を検索して行い、その推奨手法は三章で述べた2種類からオントロジ編集者が選べるように実装した。

4.2 OntRecR の評価

実装した OntRecR の評価のために、すでに研究室で行った CONSENT の運用実験データを用いた。このデータベースには、実際にエンドユーザが行動した履歴から、290組の物体間の関係と20,204件の物体操作履歴が蓄積されている。加えて物体オントロジの全要素数は3,603個、行動オントロジの全要素数は17個だった。また提案手法の評価の為に、290組の新たな関係オントロジの候補要素をオントロジ編集者の手によって作成しこれを正解セットとする。評価方法としては Recall/Precision を用いる。

RBO 法では10分割交差検定を用いて、手法の評価を行なった。まず重みを $k=0.5$ とし、レベルごとに閾値を変化させ Recall/Precision のグラフとした。その結果を図4に示す。推奨をしない場合にはすべての行動オントロジから選択するので Precision が低い、つまり選ぶ際の手間が大きくなる。それに対して、推奨を行うと Recall が下がるが Precision は大きく向上するので関係オントロジの選択の手間が削減され、関係オントロジの構築コストの削減が可能になったといえる。今回の評価ではレベルを大きくし、閾値を0~0.3あたりの値にすると良い結果が得られることがわかった。

RBH 法の場合はすべての履歴を推奨に用いた。重みは自動的に生成される。その結果を図5に示す。推奨しない場

表2 κ 統計量のインスタンス

N1,N2ともに含まれる	i11
N1に含まれ、N2に含まれない	i12
N1に含まれず、N2に含まれる	i21
N1,N2ともに含まれない	i22

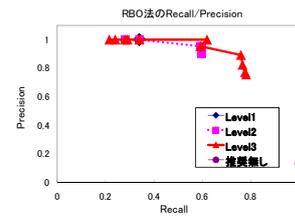


図4 RBO 法の評価

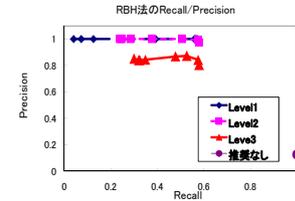


図5 RBH 法の評価

合と比べ Precision は向上したが Recall が大きく下がった。これは実際に行われた行動から推奨を行っているので、関係オントロジが存在しても、実際には行われていない行動があるということを示していると考えられる。よってこちらの手法から推奨された関係オントロジはエンドユーザの実際の行動に則しているため、この推奨結果はより空間の概念を表しているとも考えられる。RBH 法ではレベルは低く閾値でカットしない場合が最も良い結果が得られることが分かった。

5 まとめ

本論文では関係オントロジの構築の際のコストを削減するためのフレームワークと手法を提案し、その実現ツールである OntRecR を概念共有環境 CONSENT に実装し、その評価を行った。このフレームワークにより関係オントロジの構築のコストが削減でき、さらにエンドユーザの行動などをもちいて推奨することでより空間の概念を反映したオントロジの構築が可能になった。

参考文献

- [1] 小川悌知, 賀来健一, 渡邊文隆, 佐渡山英史, 富井尚志, “概念共有環境 CONSENT における行動パターンの取得”, データ工学ワークショップ論文集 (DEWS2006), 4B-i9, Mar, 2006
- [2] protégé Project, <http://protege.stanford.edu/index.html>
- [3] 溝口 理一郎, 池田 満, “オントロジー工学序説—内容指向研究の基盤技術と理論の確立を目指して”, 人工知能学会誌, Vol. 12, No. 4, pp. 559-569, 1997.
- [4] OWL Reference, <http://www.w3.org/TR/owl-ref/>
- [5] 高梨 勝敏, 佐藤 俊也, 原島 一郎, “Web コンテンツからのオントロジーの再構成方法の提案と試作”, 人工知能学会論文誌, Vol.20 No.6, pp.417-425, Nov, 2005.
- [6] 富井尚志, “マルチメディアデータベースに基づく高度コミュニティ空間の実現”, 電子情報通信学会誌, Vol.89, No.6, pp511-517, 2006
- [7] Fleiss.J.L, “Statistical Methods for Rates and Proportions”, John Wiley&Sons(1973), 佐久間 昭詔, “係数データの統計学”, 東京大学出版会, 1975.