

ライトウェイトオントロジーのゴール指向設計方法の提案と 農業分野における知識獲得への適用と評価

M2017SE008 森泉 千尋

指導教員 青山 幹雄

1 研究背景

近年、様々な分野において、AI や IoT の導入と実践により、これらを活用した自動化システムが注目を集めている。高品質な自動化システムの導入に当たり、対象領域に関する知識獲得、知識体系化の必要性がますます高まっている。システム開発における知識体系化方法の一つとしてライトウェイトオントロジーが利用されている。従来のライトウェイトオントロジー構築方法では、段階的であるが、その記述方法は確立されていない。そのためオントロジー内に不要な記述が存在する、抜け漏れの可能性があるなどの問題がある[2, 9]。

2 研究課題

本稿では、以下の2点を研究課題とする。

(1) ゴール指向に基づくライトウェイトオントロジー設計方法の提案

- オントロジーの記述範囲の設定方法の提案
- オントロジー知識拡張の制御方法の提案

(2) 提案方法の農業ドメインへの適用と評価

2.1 オントロジー

オントロジーとは情報科学において、対象世界、知識領域をある視点から見たときに現れる構成要素とそれらの関係を体系的に記述したものを指す。対象領域における構成要素がどのような意味を持つかという概念定義と階層構造の定義により、対象領域を満たす要素やそれらの関係を明確化できる[6, 7]。

2.2 ライトウェイトオントロジー

ライトウェイトオントロジーとは、オントロジーの分類の一つで以下の特徴を持つ[8]。

- 簡単な階層構造から構成される
- 対象となる知識領域が一定の範囲に限定される
- 段階的な作成、改良が可能

2.3 ライフサイクルプロセスに基づくライトウェイトオントロジー設計方法

ソフトウェア工学の技術を応用したオントロジー構築法が提案されている[2]。ライフサイクルプロセスや特性要因図などソフトウェア工学の技術をオントロジーの構築プロセスに適用し、オントロジー構築に必要な情報を段階的に詳細化し、オントロジーの記述に利用している。提案方法により対象領域に関する知識が少ないソフトウェア開発者もオントロジー構築が可能となっているが、オントロジーに不要な記述や、ぬけもれに關しての分析が行えないという問題がある。

2.4 UPON right

UPON right 法とは、UPON(Unified Process for Ontology Building)法をもとに提案されたライトウェイトオントロジー構築法[9]である。ビジネス専門家のようなオントロジーの非専門家に向けて提案されている。オントロジーの属性定義や対象領域の記述方法が確立されておらずオントロジーの記述範囲の設定が困難という問題がある。

3 アプローチ

本稿のアプローチを図1に示す。本稿ではゴール指向要求分析をオントロジー構築における記述範囲の設定に利用する[5]。トップゴールをオントロジーが示すべき知識、またオントロジーから獲得したい知識であると定義する。そのゴールを分解、詳細化した先に抽出できるタスクをオントロジーが満たすべき要求とすることで、オントロジーの記述範囲やオントロジーに記述されるべき情報をあらかじめゴールモデルと、オントロジーメタモデルに定義可能となる。作成されたオントロジーメタモデルとゴールモデルに基づいてオントロジーの構築を行う。

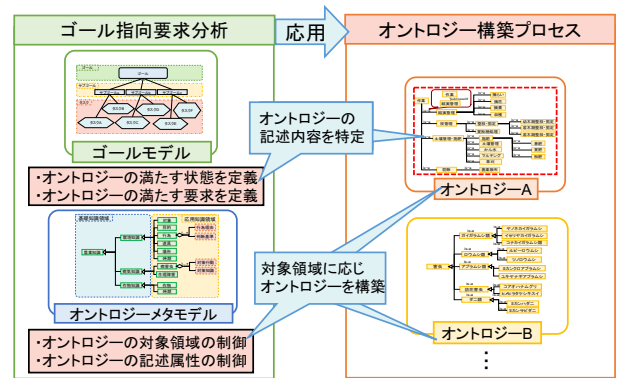


図1 アプローチ

4 提案方法

4.1 概念定義

以下に本稿における用語の概念定義を示す。

A) 基本知識領域

ある対象領域において基礎的な知識となる概念や属性を定義した知識領域。設定するゴールや作成するオントロジーによって定義される概念と属性は異なる。

B) 応用知識領域

対象領域において応用知識となる概念や属性を定義した知識領域。設定するゴールや作成するオントロジーによって追加される概念及び属性は異なる。

C) 基本オントロジー

対象領域の基礎知識領域に属する概念や属性が体系化されたオントロジー。基本オントロジーでは対象領域の基本知識が獲得できる程度のものであり応用知識領域の概念や属性は含まれない。

D) 応用オントロジー

対象領域の基礎知識領域及び応用知識領域に属する概念や属性が体系化されたオントロジー。

E) 評価タスク

オントロジーが満たすべき要求、オントロジーの記述範囲や記述されるべき情報を示したものの、知識体系の深さを評価タスクによって制御する。オントロジーの評価に用いる。

4.2 ライトウェイトオントロジーメタモデル

本稿におけるライトウェイトオントロジーメタモデルの役目は、オントロジー記述における対象領域と、対象領域に存在する属性を制御することである。オントロジーはゴール分析により作成される。図2に本稿での農業分野でのライトウェイトオントロジーのメタモデル例を示す。

図2のメタモデルによると最上位概念に農業という概念が存在し、農業知識の下位概念としては栽培知識、病気知識、作物知識が存在する。栽培知識を始めとした各概念は「対象」など複数の属性が定義されている。本稿におけるライトウェイトオントロジーはこのメタモデルにより記述する属性と対象範囲を制御する。メタモデルによると農業知識は大きく「栽培知識」、「病気知識」、「作物知識」の3領域に分類可能となる。また各領域における知識は定義されている属性を付与することができる概念のみ記述することができる。

例えば「病害虫」という属性を付与できるような概念を「栽培知識」領域のオントロジーに記述したい場合、病害虫という属性は栽培知識領域に定義されておらず属性を満たすことができないため、オントロジーの記述に表現することができない。オントロジーの記述に表現したい場合、栽培知識の領域に新しく属性を追加する必要がある。

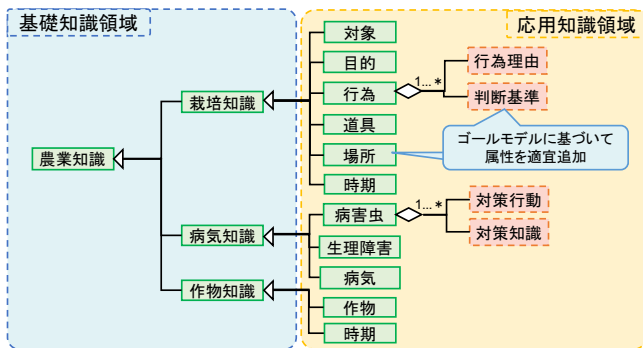


図2 農業知識におけるオントロジーのメタモデル例

4.3 提案方法

図3に本稿で提案するオントロジー構築方法を示す。以下に詳細を記述する。

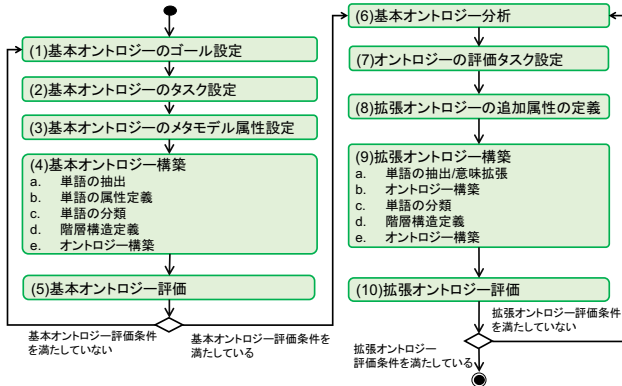


図3 提案プロセス

4.3.1. オントロジーのゴール設定

基本オントロジーのゴール設定を行う。オントロジーを作成する上で、構築されるオントロジーが満たすべき状態、オントロジーを作成する上での目標を設定する。ゴール作成後、設定したゴールを満たすためのサブゴールを設定する。サブゴールによって知識の幅を制御する。

4.3.2. オントロジーのタスク設定

サブゴールにおける目標となるタスクを設定する。設定したタスクにより構築されるオントロジーの知識の深さを制御する。図4にゴールモデルとオントロジーの関係を示す。

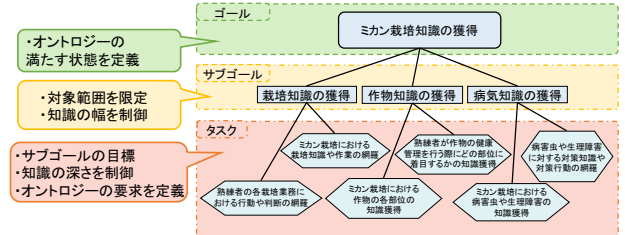


図4 ゴールモデルとオントロジーの関係

4.3.3. 基本オントロジーのメタモデル属性定義

設定したタスクをもとにオントロジー構築に用いる属性を定義する。この段階で定義された属性は基本知識領域の属性として定義される。定義される属性は、対象領域や設定するゴールによって様々である。

4.3.4. 基本オントロジーの構築

基本オントロジーの構築を行う。この時点で作成されるオントロジーは基礎知識領域の範囲の知識が獲得できる程度のものである。詳細を以下に述べる。

A) 単語の抽出

対象領域に関する資料からオントロジー構築に利用可能である単語の抽出を行う。この段階では類義語として考えられる単語であっても対象領域に関連しているものであればすべて抽出する。単語の抽出と同時に単語の意味定義を行う。対象領域において抽出された単語が対象領域において意味する内容を定義する。

B) 単語の属性定義

抽出した単語の属性定義を行う。定義する属性はライトウェイトオントロジーのメタモデルの基本知識領域内で定義されている概念の中から選択する。

C) 単語の分類

設定した属性に基づいて単語の分類を行う。このプロセスで今までに抽出された類義語をひとまとまりにする。

D) 階層構造定義

単語感の階層構造の定義を行う。単語間の is-a 関係, has-part 関係, attribute-of 関係の定義を行う。

E) オントロジー構築

A~D のプロセスによって得られた情報から基本オントロジーの構築を行う。この段階で構築されるオントロジーは対象領域の基本的知識を獲得できる程度のものであり、応用知識領域の概念は反映されていない

4.3.5. 基本オントロジーの評価

基本オントロジーの評価を行う。図5に基本オントロジーの評価の図を示す。基本オントロジーが満たすべき必要条件は以下の2点である。

A) 基本オントロジーと基礎知識領域の概念を対比し、基本オントロジーの概念が基礎知識領域の属性をきちんと満たしているか

B) 基礎知識領域と基本オントロジーを対比し、概念に矛盾がないかどうか

以上の2点の条件を満たしている場合、基本オントロジーの構築は完了したものと評価する。評価の結果基本オントロジーの評価条件が満たせていないと判断された場合、条件を満た

すことが確認できるまで (1)のプロセスに戻り、再度基本オントロジーの設計と構築を行う。

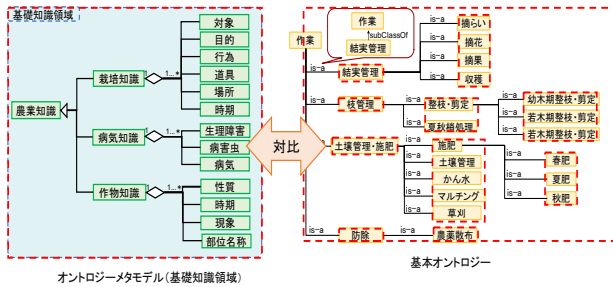


図5 基本オントロジーの評価

4.3.6. オントロジー分析

このプロセスでは基本オントロジーとゴールを分析し、現段階で構築されたオントロジーにどのような情報が不足しているか、どのような情報を付加すればオントロジーがゴールを達成できるか分析を行う。

4.3.7. 拡張オントロジーの評価タスク設定

オントロジー分析により抽出した情報をもとに、基本オントロジー構築時に抽出した、ゴール、サブゴール、タスクをさらに詳細化し、オントロジーに記述されるべき情報、獲得したい知識となる情報である評価タスクを設定する。

4.3.8. 応用知識領域の属性定義

設定した評価タスクをもとにオントロジー拡張に用いる属性の定義を行う。この段階で定義された属性は応用知識領域の属性ものとして扱う。

4.3.9. 拡張オントロジーの構築

A) 単語の抽出/意味拡張

オントロジー分析により得た情報をもとに対象領域に関する資料からオントロジー拡張に必要な単語の抽出を行う。基本オントロジー構築時に抽出した情報が応用オントロジーに利用できると判断できる場合、その情報を利用してよい。また基本オントロジー構築時に抽出しているが、意味定義の情報が足りない場合や、抜け漏れている情報が発見できた場合、この段階で適宜情報を追加する。

B) 単語の属性定義

抽出した単語の属性定義を行う。定義する属性はライトウェイトオントロジーのメタモデルの基本知識領域内、または応用知識領域で定義されている属性の中から選択する。

C) 単語の分類

設定した属性に基づいて単語の分類を行う。

D) 階層構造定義

単語間の階層構造の定義を行う。単語間の is-a 関係, has-part 関係, attribute-of 関係の定義を行う。

E) 拡張オントロジーの構築

A~D のプロセスによって得られた情報から拡張オントロジーの構築を行う。この段階で構築されるオントロジーは対象領域の基本知識及び応用知識を獲得できるものであり、基本知識領域と応用知識領域の概念がどちらも反映されている

4.3.10. 拡張オントロジーの評価

拡張オントロジーの評価を行う。図6に拡張オントロジーの評価の図を示す。拡張オントロジーが満たすべき条件は以下の3点である。

A) 拡張オントロジーの概念が基本知識領域及び応用知識領域の属性を満たしているか

B) 基本知識領域と応用知識領域と拡張オントロジーを比べ概念に矛盾がないか

C) 設定したゴール、サブゴール、評価タスクを全て達成しているかどうか

以上の3点の条件を満たしている場合、拡張オントロジーの構築は完了したものと評価する。評価の結果拡張オントロジーの評価条件が満たせていないと判断された場合、条件を満たすことが確認できるまで (6)のオントロジー分析プロセスに戻り、再度拡張オントロジーの設計と構築を行う。

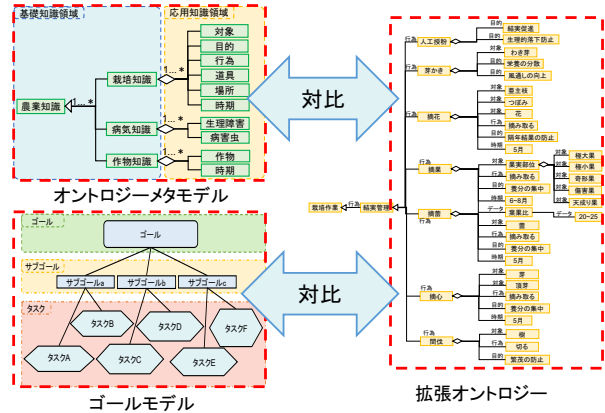


図6 拡張オントロジーの評価

5 提案方法の例題への適用

5.1 農業分野への適用

本稿では、提案プロセスを農業分野の知識獲得に用いることにより、提案方法の妥当性および有用性の評価を行う。農業分野は一般的に知識の獲得に多くの時間を要する。また、農業分野には様々な要素が存在しており、知識体系化が非常に重要となる分野であるといえる。したがって、本稿の適用例題として農業分野は適していると考えられる。

5.2 前提条件

例題の適用にあたり以下の前提条件を設定する。

- (1) 本稿では、例題適用における対象領域を農業分野のミカン栽培に関する知識獲得とする
- (2) ゴール、サブゴール設定、タスク及び評価タスク設定は正しいものが設定されているとする。

5.3 例題におけるオントロジーの構築

提案プロセスを農業分野のミカン栽培における知識獲得に適用した。知識獲得は資料から行った[1, 3, 4, 10]。表1に例題で構築したオントロジーの全概念数の集計結果を示す。農業ドメインにおける例題適用ではゴール分析とオントロジーメタモデルの構築により、「栽培作業知識」、「病気知識」、「病害虫知識」、「生理障害知識」の4領域のオントロジーが作成された。オントロジーの概念数を領域ごとに表2-5に示す。

表1 オントロジー全概念数

知識領域	基本知識領域	応用知識領域
概念数	58	309

表2 栽培知識領域のオントロジー概念数

属性名	行為	対象	目的	時期	道具	データ	作業	合計
概念数	29	43	28	10	7	5	28	150

表3 病気知識領域のオントロジー概念数

属性名	病気	部位	症状	影響	合計
概念数	12	24	40	1	77

表4 病害虫知識のオントロジー概念数

属性名	病害虫	部位	頻度	影響	時期	合計
概念数	19	23	8	51	5	106

表5 生理障害知識のオントロジー概念数

属性名	生理障害	部位	症状	合計
概念数	9	9	21	39

6 例題におけるオントロジー評価

6.1 ゴール分析による記述範囲の設定

農業ドメインにおいて栽培に関する概念、また栽培において主要となる情報を特定する「目的」や「対象」といった属性が特定された。また病気や病害虫に関して「症状」や「影響」といった病害虫の情報を詳細に定義する属性が特定された。ゴール指向要求分析をオントロジー構築プロセスに適用し、オントロジーの記述範囲や属性をメタモデルとゴールモデルに定義することで、オントロジーに記述されている知識の重要度が分析可能となった。

6.2 オントロジー拡張のための情報分析

表 6-8 において病気知識の領域では「病気による影響」を表す属性を持つ概念が少ないことが分かる。また、表 6-9 の生理障害知識のオントロジーには「生理障害による影響」に関しての属性設定や概念の記述が存在していないことが分かる。より正確なオントロジーを構築するにはこれらの情報の詳細化が必要と分析できる。本提案方法によりオントロジーの概念分析が可能となり、拡張に必要な情報を収集可能になることを示すことができた。

6.3 オントロジー記述内容の分析

例題では、オントロジーの記述に必要なが、属性定義されていなかった概念が 5 つ見つかった。提案方法によりゴールモデル、オントロジーメタモデルとオントロジーの記述を比較できるようになり、オントロジー記述内容に抜け漏れている情報やオントロジーに不要な記述を分析可能になった。

7 評価

7.1 オントロジー構築における知識の制御

本稿ではゴール要求分析指向により対象領域の定義と、対象領域において記述すべき概念や情報をあらかじめ評価タスクやオントロジーメタモデルに定義可能となった。これにより以下の 2 点を行えるようになる。

(1) ゴール分析による記述範囲の制御

体系化する知識の幅をゴールで、深さをタスクで表現することにより、作成するオントロジーの記述範囲の制御をソフトウェア工学技術者などオントロジーの専門家でなくても行えるようになった。ゴール要求指向分析を用いて評価タスクやのオントロジーメタモデルの詳細化を繰り返し行うことで、段階的にオントロジーの記述範囲を設定する事ができるようになり、より構造的なオントロジーを記述方法が確立できたといえる。

(2) オントロジーの分析による知識拡張の制御

提案方法により、構築したオントロジーをメタモデルに基づいて分析可能となった。構築したオントロジーの記述をオントロジーのメタモデル、評価タスクに基づいて分析することで、構築されたオントロジーに不要である記述の発見、現時点のオントロジーに不足の情報を分析可能となった。この分析によりオントロジー作成者による知識拡張の制御が可能になった。

8 考察

8.1 過去の文献との比較

(1) オントロジー記述範囲の設定

従来のオントロジー構築法[2, 9]ではオントロジーの属性定義や対象領域の記述方法が確立されておらずオントロジーの記述範囲の設定が困難であった。本稿では、オントロジーが満たすべき状態や、オントロジーに記載するべき情報を抽出することでオントロジーメタモデルに、必要な概念や属性を設定可能とした。オントロジー構築における対象範囲を明確化可能とした記述方法を提案したという点で意義があると言える。

(2) オントロジー分析の観点

従来のオントロジー構築方法では[2, 9]オントロジーに不要な記述や、ぬけもれに関しての分析が行えなかった。本稿では、オントロジーメタモデルによりオントロジーの対象領域と属性を設定することで、構築されたオントロジーの概念を属性ごとに表現可能となり、構築されたオントロジーに不必要である記述や不足の情報を分析可能になった。この分析により段階的なオントロジーの記述拡張を行えるようになり、より構造的なオントロジーの記述方法を実現したといえる。

9 今後の課題

(1) オントロジーのゴール設定

(2) オントロジー情報補充のためのプロセス提案

(3) オントロジーの分析プロセスの再構築

10 まとめ

従来のライトウェイトオントロジー構築方法では、段階的であるが、その記述方法は確立されていない。本稿では、ゴール指向要求分析をオントロジーの記述範囲の設定に応用したオントロジー構築方法を提案した。あらかじめオントロジーに必要な情報や、オントロジーの満たすべき要求を抽出することでオントロジーの対象領域と存在する属性を定義したメタモデルの構築が可能となった。さらに、ゴールとメタモデルに基づくオントロジー構築により、オントロジー段階的拡張のための分析や、オントロジーの対象領域と記述範囲の制御を可能とした。提案方法を農業分野におけるミカン栽培の知識獲得に適用し提案方法の有効性を評価した。

参考文献

- [1] 有江 力, 図解でよく分かる病害虫の基本 病害虫発生メカニズムから、栽培管理、農薬・肥料の使い方、防除法まで, 誠文堂新光社, 2016
- [2] R. Fujimoto, et al., A Lifecycle-Based Design Methodology of Lightweight Ontology and its Application to Cultivating High Quality Mandarin Orange, Proc. of APSEC 2014, IEEE, Dec. 2014, pp. 154-157.
- [3] 堀江 武, 図解でよく分かる農業の基本 栽培の基礎から新技術, 流通, 収納まで, 誠文堂新光社, 2015.
- [4] 岸野 功, ミカンの作業便利帳, 農山漁村文化協会, 2012.
- [5] JISAREBOK 企画 WG, REBOK, 近代科学社, 2011.
- [6] 溝口 理一郎, オントロジー構築入門, オーム社, 2006.
- [7] 溝口 理一郎, オントロジー工学, オーム社, 2006.
- [8] 溝口 理一郎 他, web2.0時代のオントロジー利用雑感 ライトウェイトからヘビーウェイトまで, 人工知能学会研究会, 人工知能学会研究会資料, SIG-SWO-A602-06, 2006, pp. 1-8.
- [9] A. D. Nicola et al., A lightweight Methodology for Rapid Ontology Engineering, CACM, Vol. 59, No. 12, Mar. 2016, pp. 79-86.
- [10] 杉浦 明, 果樹栽培の基礎, 農山漁村文化協会, 2013.