

多言語利用可能なオントロジーを使った遠隔教育コンテンツ検索システムの有用性評価

千葉栄一 (NTT コミュニケーションズ)
中野美知子 (早稲田大学)
小倉健太郎、河村誠司 (NTT ソフトウェア)

言語、文化の違いを意識することなく教育コンテンツを検索できるオントロジーベース検索機能の効果を持続して利用するために、オントロジー辞書の拡張を自律成長させる機能の検証を行い、実利用における運用面での有用性評価を行った。

Utility evaluation of the remote education contents search system which uses the multilingual ontology.

Eiichi Chiba (NTT Communications)
Michiko Nakano (Waseda University)
Kentaro Ogura, Seiji Kawamura (NTT Software)

Without being conscious of the difference in the language and the culture, we search for remote education contents, a possible ontology base searching function has been developed. The function self-control makes expansion of an ontology dictionary grow from which was inspected to make the effect of the searching function continue.

1.はじめに

現状、アジア各国間には顕著な情報格差があり、この格差はアジア域内の産業や文化を活性化するうえでの大きな障壁となっている。情報格差が発生する要因として、文化の違い(≒表現の違い)、言語の違い等によって、アジア域内に情報が存在するにも関わらず、必要となる教育コンテンツ(以下、L0: Learning Object)等が見つけられず、取り出せない環境が課題として挙げられる。

この課題の解決に向け、総務省 アジア・ブロードバンド計画における遠隔教育実験(平成17年度～平成19年度の3ヵ年)に

て遠隔教育コンテンツへの応用を目指したオントロジー検索技術の実証実験を行い、平成18年度実験では、言語・文化の違いを意識することなく、それぞれの母国語である日本語、タイ語でL0を検索できる、“オントロジーベース検索機能”(図1)を開発し、概念のID化による言語の違い、オントロジー辞書の拡張による文化の違いの吸収を行い、その有用性を実証した。(研究報告書: 2007-CE-91 に関連論文)

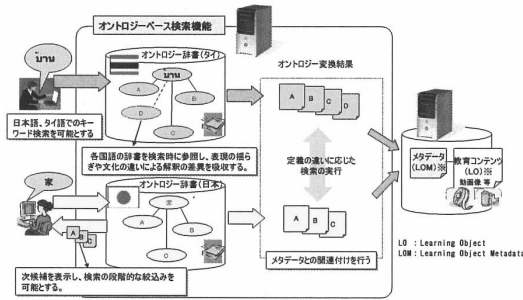


図 1 オントロジーベース検索機能

また、オントロジーベース検索機能の中核であるオントロジー辞書の変換例を図 2 に示す。日本語、タイ語のそれぞれのオントロジー辞書で“切る”という ID 化された語彙に関連する変換結果を、LOM (Learning Object Metadata) から抽出することで、それぞれの母国語による検索を可能にしている。

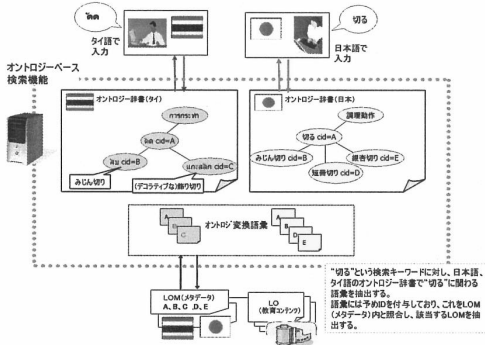


図 2 オントロジー辞書変換概略図

平成 19 年度実験では、オントロジーベース検索機能の効果を持続して利用するため、オントロジー辞書を自律成長させる機能の検証を行い、実利用における運用面での有効性評価を行った。

2. オントロジー辞書の自律成長における考慮点

オントロジー辞書の自律成長機能のシステム実装において、下記 a, b の 2 点を満た

す方式として L0 の中心概念間のみを管理する、“コンテンツ適応型ドメインオントロジー方式”を考案した。(表 1)

以下に実装への考慮点を示す。

a. L0 検索者に不利益が生じないこと

考慮点 1：オントロジー辞書自律成長の実装による検索精度低下への対策

→ L0 は、特に大学等の高等教育分野では専門性が要求され、一般的な体系でのオントロジー辞書（例：食のドメイン）は不向きである。よって、オントロジー定義の登録には専門家の知識が必要であり、L0 内容に熟知した L0 登録者が検索に必要十分と思われるオントロジー定義を記述することで、検索精度低下への対策を講じた。

考慮点 2：L0 登録を公開することに対する不正な登録への対策

→ L0 登録インターフェースを公開し、オントロジー辞書に対し概念の追加を行うと、不適切なオントロジー辞書に成長するリスクが発生する。このリスク対策として、L0 登録者には自由にオントロジー辞書を編集させずに、自身が登録する L0 のオントロジー定義のみを編集させた。

→ L0 登録者のオントロジー辞書編集への制約を設けたことで、L0 間の関係を整備する“サービス管理者”が必要となった。

b. 現実的な運用サイクルが設計されていること

考慮点1：L0登録者が他L0との関係も考慮してオントロジー定義を行うことは困難

→ L0登録者に他L0との関係定義を行わせることは、興味のないL0との関係定義をさせることになり、作業実施モチベーションの維持が難しい。その結果、「L0間の関係が定義されない」「不適切なL0間の関係が定義される」など、検索者の不利益につながる可能性がある。
(本観点からも「a. L0検索者に不利益が生じないこと 考慮点2」の方針は適切と判断できる。)

考慮点2：サービス管理者が、全てのL0の中身を理解して、適切なオントロジー定義を行うことは困難

→ 「a. L0検索者に不利益が生じないこと 考慮点2」により、サービス管理者がL0間の関係記述およびオントロジーの調整を行う必要が発生する。しかし、全てのオントロジーの調整を行うためには、「全オントロジー調整の作業コスト $\approx L \times n$ (L=L0数, $n=L0$ に付与される平均キーワード数)」の作業コストが発生する。この問題を回避するために、「L0の最も代表的な概念を表す“中心概念”を設定し、中心概念間の関係のみを管理する“コンテンツ適応型ドメインオントロジー方式”を考案した。これにより、作業コストをL0数に抑え

ることが可能となった。

表1 コンテンツ適応型ドメインオントロジー方式

観点	一般的なオントロジー辞書 (従来方式, H18年度)	コンテンツ適応型ドメインオントロジー辞書 (新方式, H19年度)
	オントロジー辞書 構成	
オントロジー辞書 成長方式	該当する語彙をtree状に手動更新する。	L0登録時の中心概念から、中心概念同士の繋がりを手動更新する。
検索方式	シングルパス検索	マルチパス検索 (L0検索者が検索した文字列に従って、複数の該当する同一語彙から情報を取り出す)
オントロジー定義 の難易度	難	易
作業コスト	$L \times n$	L

このコンテンツ適応型ドメインオントロジー方式によるオントロジー辞書は、L0登録→中心概念の自動更新→中心概念同士の自動更新による自律成長サイクルとなる。

(図3)

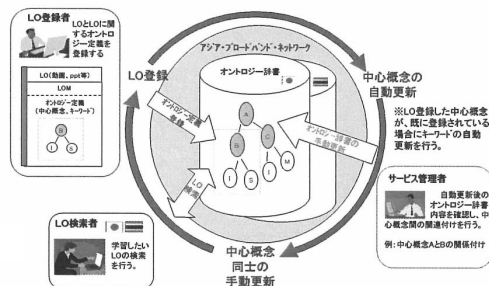


図3 オントロジー辞書の自律成長サイクル

3.実証実験内容と結果

(1) 実験内容

日本の早稲田大学、及びタイのチュロンコン大学を教育機関実験フィールドとして、L0登録者、L0検索者、サービス管理者の役割を設定したうえで、オントロジー検索の自律成長の有用性を評価するために下記6点(実験①～実験⑥)を評価項目として実験を行った。(図4)(表2)

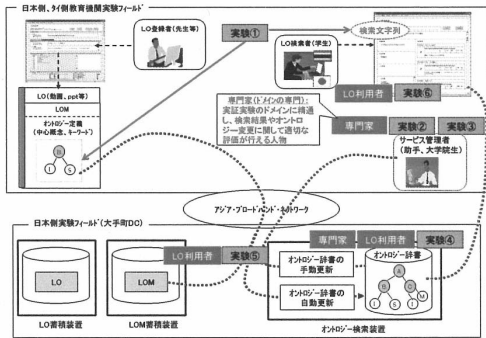
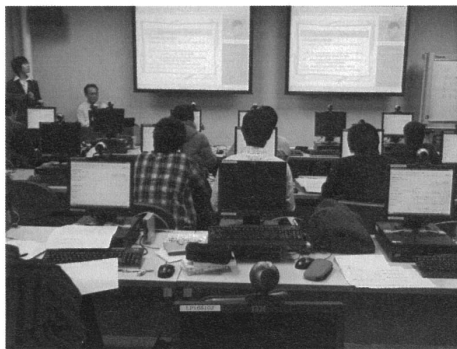


図4 オントロジー辞書の自律成長の有用性実験イメージ

表2 実験評価項目と観点

実験番号	実験評価項目	評価観点
オントロジー辞書の自律成長の有用性	① タギングされたオントロジー定義のLOM検索での活用性	LO登録者が適切なオントロジー定義を記述できるかを評価する。
	② オントロジー辞書自律更新の正確性	サービス管理者が適切にオントロジー辞書上で中心概念間の関係付けができるかを評価する。
	③ オントロジー辞書手動更新の必要性	サービス管理者が行うオントロジー辞書の手動更新がLO検索に寄与していることを評価する。
	④ オントロジー辞書の自律成長サイクルの有用性	自律成長でより有効な検索結果得られるかを評価する。
	⑤ オントロジー辞書自律成長下の、手動更新前後でのLOM検索の利便性	サービス管理者が行うオントロジー辞書の手動更新がLO検索に寄与していることを評価する。
	⑥ オントロジー辞書自律成長下の、段階的な検索インターフェースの利便性	上位/下位、同義、関連の関係を使った検索インターフェースは使い易いか評価する。



実験風景 (早稲田大学)

(2) 実験結果

実験結果を表3に示す。この結果から、オントロジー辞書の拡張・自律成長のために重要な役割を担うオントロジー定義の中心概念の抽出、中心概念相互の関係付け、マルチパス検索などの仕組みが有

効に機能することを確認し、コンテンツ適応型ドメインオントロジー辞書方式におけるオントロジー辞書の自律成長の仕組みがLO検索基盤として有用であることが検証できた。

表3 実験結果概要

実験番号	実験結果概要	結論
①	LO登録時のオントロジー定義とLO検索者の検索キーワードの一致度(表現の揺らぎ含む)は42%であった。(N:444回)	LO検索者の検索キーワードは、オントロジー定義とほぼ一致し、目的とするLOを検索できる。
②	サービス管理者の中心概念間の関係付けを専門家が適切とした評価した割合は100%であった。(N:16)	サービス管理者は適切にオントロジー定義の中心概念間の関係付けができる。また、LO検索者の検索効率向上にも寄与する。
③	手動更新により新たにLO検索できた専門家の平均評価は4.6点/5点満点であった。	
⑤	オントロジー辞書の自律成長サイクル後のLO検索者の利便性平均評価は3.9点/5点満点、専門家の評価は5点満点であった。	
④	オントロジー辞書の自律成長サイクル後のLO検索者の利便性平均評価は3.9点/5点満点、専門家の評価は5点満点であった。	オントロジー辞書の自律成長サイクル後は、自律成長前に比べLO検索者にとってより有効なLO検索ができる。
⑥	LO検索者の100%が「とても便利」、「便利」との評価であった。	検索インターフェースとして、「上位」「下位」「関連」の機軸が表示されることはLO検索の利便性向上に寄与する。

4.実験終了後の期待効果

本実証実験で開発したオントロジーベース検索機能をOSS (Open Source Software) とし、各教育機関で遠隔教育におけるLO検索基盤として運用することで、アジア域内へ日本のLOを発信することも可能となる。
※LOの公開が前提

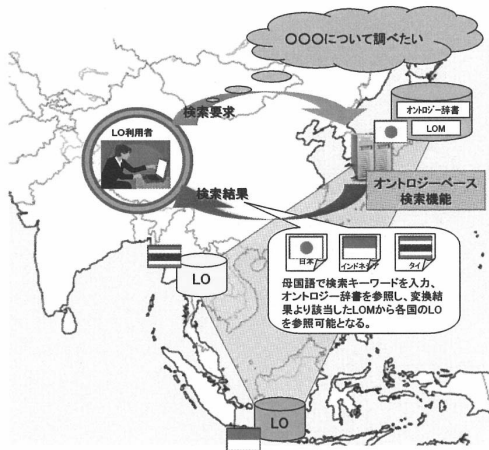


図5 アジア域内での利用シーン

謝辞：本文は、平成 19 年度の総務省プロジェクト「国際情報通信ハブ形成のための高度 ICT 共同実験」の一環として実施されたものである。