

コンテンツ検索のためのオントロジーの構築と e-ラーニングシステムへの実装

砂原 悟[†] 友田 成則[†] 今井 順一[†] 小松川 浩[†]

千歳科学技術大学[†]

1. はじめに

昨今、インターネットを利用した学習に対する機運の高まりや、通信環境の向上によりインターネット上に多種多様の学習コンテンツが分散して存在する。利用者は必要なコンテンツを検索機能で探すことになるが、日本語では主にひらがな、カタカナ、漢字、英字を使用しており、同義であるが異なる語彙でコンテンツが登録されていることも少なくないため、必要なコンテンツ探し出すことが困難になっている[1]。この問題を解決するために我々は知識システムを構成するときに用いる基本概念および語彙を体系的に表すオントロジーに着目し、オントロジーの中でも語と語の意味関係や知識関係を定義するシソーラスを採用した。我々はこれを理工系大学の初等数学の範囲で構築し、検索機能に適用することで語彙の差(表記揺れ)を吸収することに成功した。オントロジーの構築には数学教員の知識とユーザの検索キーワードに着目し、相互の関係と別称を定義した。次章より詳しく説明する。

2. オントロジーと e-ラーニングの関係

シソーラスは単に語彙の差を吸収するだけでなく、is-a, has-a の関係が明確に示せるため、学習手順のひとつの指標になると考えられる[2]。例えば数学において微分の定義に極限の概念を含むため微分 has-a 極限と定義し、微分を習得するために極限が必要であると示すことができる。

オントロジーは現在 Web の要素技術として W3C で標準化されており、Web アプリケーションである e-ラーニングシステムへの組み込みも容易であるなど有意義な点が非常に多い。

その事実をふまえ、我々はオントロジーエディターである Protégé [3]を用いて約 80 の語彙

を定義した。語彙の一部を図 1 に示す。我々が構築したオントロジーは OWL Lite に準拠しており、我々が保有しているシステムの構造からカリキュラム、カテゴリ、キーワードをクラスとして定義し、インスタンスを表 1 のように構築した。またインスタンス同士の関連を表す語彙として Dublin Core[4]とその精密要素を採用した(表 2)。各インスタンスには数学教員の知識を基に is-a, has-a の関係を定義した。また、e-ラーニングの検索ログを参照しながら別称を定義した。

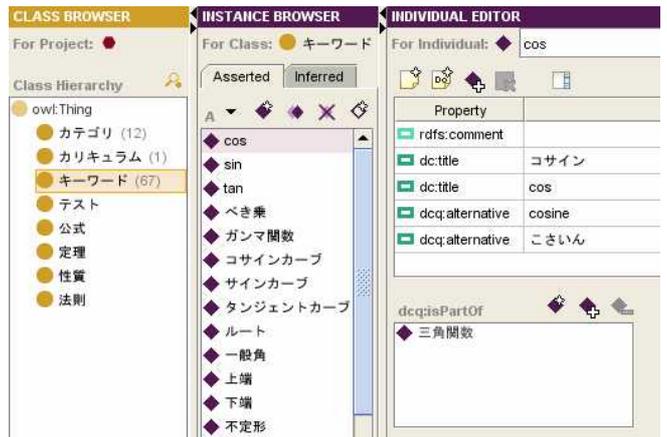


図 1 構築したオントロジー

表 1 オントロジーの構成

クラス	インスタンス
カリキュラム	数学
カテゴリ	三角関数 不定積分 弧度法 極限など
キーワード	Sin 正接 周期 積分定数など

Implementation of Retrieval Function in Ontology-based learning System
Satoru Sunahara, Masanori Tomoda, Junichi Imai, Hiroshi Komatsugawa
Chitose Institute of Science and Technology

表 2 関係を表す語彙

dc : relation	リソースの参照 例 : 積分 微分
dcq : isPartOf	is - a の関係 例 : 同次形 微分方程式
dcq : hasPart	Has - a の関係 例 : 微分方程式 同次形
dcq : alternative	別名であるが同一の リソースの関係 例 : tangent tan

dc : Dublin Core
dcq : Dublin Core 精密要素

3. e-ラーニングへの適用

我々がこれまでに開発してきた e-ラーニングシステムには、数学・物理・化学など 10000 を超えるコンテンツが存在する。特に数学においては、小学・中学・高校・理工系大学の基礎までを体系的に整備するに至っている。コンテンツは SCORM モデルを採用しており、コンテンツ単体を Asset と定義し、また Asset に対してメタ情報を記述する LOM も定義した。さらに節単位で Asset をまとめたものを SCO とし、SCO 同士の is - a, has - a の関係も独自に知識 LOM として定義した。また LOM を用いたコンテンツ検索機能が実装されており、SCO の名前、LOM のタイトルおよびキーワード、知識 LOM のキーワードを対象としたキーワードマッチングによる検索を行うものとなっている。

検索機能へオントロジーの適用をするにあたり、is - a, has - a の関係をオントロジーで定義し、知識 LOM のキーワードを LOM のキーワードへ移行した。システム上でオントロジーを扱うために Jena フレームワーク [5] と RDF・オントロジーのクエリ言語である SPARQL [6] を組み込んだ。以下に実装した検索の大まかな流れを示す。

- 1 ユーザが入力した検索キーワードに関連する語彙がオントロジーにあれば、それを検索キーワードに加える。
- 2 コンテンツの LOM に一致するものがあれば、検索結果として加える。
- 3 検索キーワードが複数 AND 条件で絞り込む。

4. 結果

検索機能にオントロジーを適用することにより、表記揺れを吸収することが可能となった(図 2)。また、検索結果が飽和した場合、is - a, has - a の関係を考慮したキーワードを絞り込みのヒントとして提示することが可能となった。

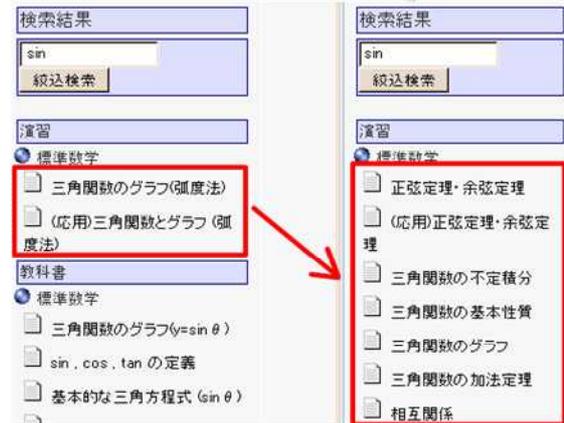


図 2 検索結果

5. まとめ

オントロジーを用いることにより、教員と学習者の間における語彙の差(表記揺れ)を吸収し、よりユーザがコンテンツを見つけやすい環境を構築することができた。

今後の課題としてオントロジーとして定義されている語彙が約 80 と非常に少なく、現在は数学に限られているため、コンテンツが存在している各分野において語彙を拡張していく必要がある。また、現在の検索結果の表示順序はデータベースに依存した順序であるため、参照率の高いコンテンツを優先的に表示するなどの検索ロジックに対する工夫などが必要である。

参考文献

- [1] Satoru Sunahara, Hiroshi Komastugawa etc. : Ontology based Learning System for Science and Technology
- [2] 林雄介 溝口理一郎 他: オントロジーアウェアな学習コンテンツ設計環境
- [3] ontology editor protégé : <http://protege.stanford.edu/>
- [4] Dublin Core <http://dublincore.org/documents/dces/>
- [5] Jena Semantic Web Framework <http://jena.sourceforge.net/>
- [6] SPARQL Query Language for RDF <http://www.w3.org/TR/rdf-sparql-query/>