

토픽 맵을 이용한 지능형 학습 콘텐츠 구축 방안

이호영*

요약

인터넷 환경의 발전과 멀티미디어 콘텐츠의 데이터가 대용량화되고 있는 추세이다. 특히 교육 콘텐츠의 개발과 활용이 일반화되고 있으며, 사용자가 원하는 정보를 검색하는데 많은 문제들이 발생하고 있다. 기존 대부분의 이러닝(E-Learning) 사이트의 학습 콘텐츠 검색 시 단순히 제목과 같은 텍스트 위주의 검색만 가능하기에 그 내용적 검색이 힘들다는 문제점을 가지고 있다. 이에 본 논문에서는 시맨틱 웹의 핵심인 온톨로지(Ontology)를 사용하여 학습 콘텐츠의 메타데이터를 구성하고 학습자에게 이를 기반으로 의미적인 이러닝 검색 시스템을 지원하고 협력 필터링 기법을 사용하여 효율적인 학습 콘텐츠 검색이 가능하도록 하였다.

The Method of Intelligent Learning Contents Building based on Topic Map

Ho-Young Lee *

ABSTRACT

Lately, It is trend that development of internet environment and data of multimedia contents are done high-capacity. Specially, development and practical use of education contents are generalized, and much problems are happening searching information that user wants. We have controversial point that search of contents is hard because was possible text search such as title simply when search studying contents of most E-Learning site in existing. In this paper, we compose MetaData of studying contents using Ontology that is point of semantic web. Also, by we support semantic E-learning search system to learner and use cooperation filtering techniques, did so that efficient studying contents search may be available.

Key Words : contents, E-Learning, MetaData, Ontology, semantic

* 초당대학교 정보통신공학과(✉hylee@chodang.ac.kr)

· 제1저자(First Author) : 이호영 · 교신저자(Correspondent Author) : 이호영

· 접수일(2011년 4월 7일), 수정일(1차 : 2011년 5월 13일), 게재확정일(2011년 5월 17일)

I. 서론

인터넷 환경의 발전과 멀티미디어 콘텐츠의 데이터가 대용량화되고 있는 추세이다. 특히 교육 콘텐츠의 개발과 활용이 일반화되고 있으며, 사용자가 원하는 정보를 검색하는데 많은 문제들이 발생하고 있다. 기존 대부분의 이러닝(E-Learning) 사이트의 학습 콘텐츠 검색 시 단순히 제목과 같은 텍스트 위주의 검색만 가능하기에 그 내용적 검색이 힘들다는 문제점을 가지고 있다. 또한 학습자의 요구에 맞는 코스웨어(Courseware) 추천의 필요성이 인식되고 있으며 이에 맞는 웹기반 교육의 효율적이고 자동화된 시스템의 필요성이 대두되고 있다.

위와 같은 문제점을 보완하기 위해서 기존의 키워드 중심의 검색 방식에서 벗어나 컴퓨터가 이해할 수 있도록 정보에 의미를 부여하여 빠르고 정확한 검색을 할 수 있도록 시맨틱 웹(Semantic Web) 기술이 등장하였는데, 시맨틱 웹은 현재의 컴퓨터처럼 사람이 마우스나 키보드를 이용해 원하는 정보를 찾아 눈으로 보고 이해하는 웹이 아니라, 컴퓨터가 이해할 수 있는 웹을 말한다. 즉 사람이 읽고 해석하기에 편리하게 설계되어 있는 현재의 웹 대신에 컴퓨터가 이해할 수 있는 형태의 새로운 언어로 표현해 기계들끼리 서로 의사소통을 할 수 있는 지능형 웹이다. 이는 사람들이 이해할 수 있도록 자연어 위주로 되어 있는 현재의 웹 문서와 달리, 정보 자원들 사이에 연결되어 있는 의미를 컴퓨터가 이해할 수 있는 형태의 언어로 바꾸는 것이다. 이렇게 되면 컴퓨터가 정보 자원의 뜻을 해석하고, 기계들끼리 서로 정보를 주고받으면서 자체적으로 필요한 일을 처리하는 것이 가능해진다.

이에 본 논문에서는 시맨틱 웹의 핵심인 온톨로지(Ontology)를 사용하여 학습 콘텐츠의 메타데이터를 구성하고 학습자에게 이를 기반으로 의미적인 이러닝 검색 시스템을 지원하고 협력 필터링 기법을 사용하여 효율적인 학습 콘텐츠 검색이 가능하도록 하였다.

II. 관련 연구

토픽 맵은 개념을 정의할 수 있는 객체를 토픽으로 정의하고 이를 통해 도메인을 실제적으로 명세화할 수 있다. 토픽 맵은 인텍싱, 시소러스, 목차 등의 구조를 다루기 위해 사용되며, 특정 응용 분야를 확장할 수 있을 뿐 아니라 토픽의 위치 정보를 나타내는 멀티 차원의 토픽 스페이스(Topic Space)를 정의한다. 이러한 토픽 맵을 사용하여 복잡한 정보나 단순하고 모호한 정보까지 지식표현 스키마의 전체적인 범위에서 처리할 수 있다는 장점이 있으며, 일반적인 상위 수준 개념이나 내부적 연결 정보 오브젝트 셋을 관리하기 위한 프레임워크(Framework)를 제공한다[1][2][3][4][5][6]. 또한 토픽 맵(Topic map)은 정보를 필요에 따라 연결하고 분류하여 사용자가 손쉽게 이용할 수 있게 함으로써 효과적인 정보 검색을 할 수 있도록 설계된 web 3.0의 표준기술이다[1][7]. 토픽 맵은 토픽(Topic), 어소시에이션(Association), 어커런스(Occurrence)의 개념을 사용하여 객체들 간의 연관성을 그림 1과 같이 표현한다[8][9].

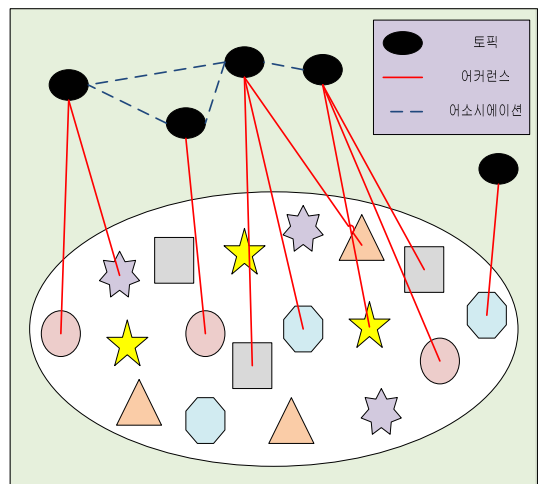


그림 1. 토픽 맵에서의 관계
Fig. 1. Relation in the Topic Map

III. 전체 시스템의 구성

지능형 학습 콘텐츠 구축은 그림 2와 같이 구성되어 있다. 이를 크게 2개의 부분으로 나누어보면 첫째, 선호도 프로파일(Preference Profile)과 트랜잭션 프로파일로 구성된 메타데이터(Metadata)로, 학습자의 나이, 선호과목, 사용언어, 선호강사 등의 정보를 XML 형태로 구성하며 이는 향후 추천시스템을 통해 학습 성향이 피드백 된다. 이를 위해 각 학습 콘텐츠의 중요도를 학습자가 직접 1 ~ 5 까지 입력하게 하여 향후 타 학습 콘텐츠의 추천 시 사용하게 된다. 뿐만 아니라 사용자의 연령대, 선호과목 등을 토대로 새로운 학습자에게 학습 콘텐츠 추천 시 이를 활용할 수 있게 한다.

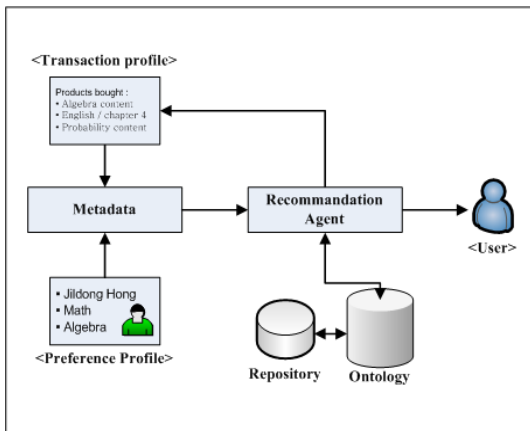


그림 2. 전체시스템 구성도
Fig. 2. Diagram of Total System

이러한 메타데이터에 대한 내용은 4장에서 자세히 기술하도록 하겠다. 두 번째는 학습 관련 정보 저장소와 이에 대한 토픽 맵 구성이다. 일반적으로 학습 정보는 SCORM 기반으로 구성되어 있으며 이는 5장에서 자세히 설명하겠으며, XML 형태로 이를 쉽게 토픽 맵으로 변환할 수 있다.

IV. 메타데이터 구성

선호도 프로파일은 사용자 프로파일을 바탕으로 만들어진 것이며, 트랜잭션 프로파일은 사용자가 인터페이스에 접속하여 입력한 검색어, 참조 횟수 등을 바탕으로 작성된다. 즉, 사용자 데이터베이스와 사용자 성향 데이터베이스를 생성하고, 이를 바탕으로 사용자의 기호를 나타내는 메타데이터를 작성하게 된다. 또한 추천 내용을 피드백 하여 사용자 선호도 프로파일을 업데이트한다. 이를 위해 협업 필터링을 사용하여 학습자의 관심 분야와 학습 객체 정보의 학습 권장자 단어의 유사도를 계산하고 강사명과 과목명을 검색한 횟수 그리고 학습을 경험한 빈도에 의해 선호도를 분석한다. 유사도와 선호도는 피어슨의 상관관계 공식[11]과 사용자의 항목에 대한 예측을 구하기 위해 GroupLens에서 제안한 식[12]을 이용하여 측정하였다. 피어슨 상관관계 방법은 두 변수 사이에 존재하는 선형관계의 정도를 측정하며, 상관계수는 선형 리그레션 모델에서 유도된다. 이러한 로그 데이터에 기록된 자료를 바탕으로 학습자 프로파일 데이터베이스에 따로 저장되게 한다. 학습자의 과목, 강사에 대한 선호도 성향은 시간에 따라 변화하므로 학습 에이전트를 사용하여 피드백의 반복적인 학습에 의한 분석이 이루어진다.

$$W_{a,k} = \frac{\sum_i (r_{a,j} - \bar{r}_a)(r_{k,j} - \bar{r}_k)}{\sqrt{\sum_i (r_{a,j} - \bar{r}_a)^2 * \sum_i (r_{k,j} - \bar{r}_k)^2}} \quad (1)$$

식(1)에서는 유사도를 나타내고 있으며 j는 사용자 a와 사용자 k의 의해 공통된 선호도를 의미하고, r_{xy} 는 사용자 x의 정보 y에 대한 선호도를 나타내며 \bar{r}_x 는 사용자 x의 전체 정보에 대한 평균 선호도이다.

$$P_{a,j} = \bar{v}_a + \frac{\sum_{i=1}^n w_{a,i} (v_{i,j} - \bar{v}_i)}{\sum_{i=1}^n w_{a,i}} \quad (2)$$

식(2)에서 선호도는 사용자 **a**의 정보 **j**에 대해서 예측한 선호도이며 \bar{v}_a 는 사용자 **a**의 평균 선호도이다. 본 논문에서는 학습자 프로파일을 이용하여 학습자의 과목, 학습자 강상진에 대한 선호도 및 성향은 시간에 따라 변화하므로 학습 에이전트를 사용하였다. 또한 피드백을 위해 반복적인 학습에 대한 분석결과를 학습자의 강의정보에 대한 재학습 빈도수의 선호도, 수업누적 시간 선호도, 과목 선호도를 메타데이터로 생성하였다.

표 1. 학습자 메타데이터 구축의 예
Table 1. Example of Meta Data Building

```

<interest>
  <interestfield>컴퓨터프로그래밍언어
</interestfield>
  <interestprof>강상진</interestprof>
  <interestcont>C#          프로그래밍          입문
</interestcont>
</interest>
...
<Learn_language>
  <TheFirst>영어</TheFirst>
  <TheSecond>일본어</TheSecond>
</Learn_language>
    
```

V. 토픽 맵을 이용한 학습 콘텐츠 구축

대부분의 SCORM 기반의 학습 활동단위 구조는 계층적으로 구성된다. 본 논문에서는 도메인으로 e-learning을 선정하고 학습과정, 학습객체 메타데이

터, 학습자 메타데이터를 클래스로 구성하였다. 정확한 정보를 연결하고 구조화 하기 위해 계층적인 구조로 구축하였다. 이를 위해 XTM(XML Topic Map)을 사용하여 구축하였는데, XTM은 토픽 맵을 표현하는 기술 언어로 토픽 맵 모델의 각 요소를 나타내는 태그 집합과 엘리먼트들 사이의 구조를 정의하고 있으며, 토픽 맵을 기술하는 표준포맷으로 사용되고 있다는 장점을 가지고 있다. 학습과정 클래스는 SCORM1.2를 기반으로 웹상에서 실행되고 있는 e-learning 사이트들의 학습 활동단위 구조를 표현하며 학습객체 메타데이터는 학습자가 학습과정에 필요한 요소들에 연결되어진 각 학습 객체들로 이루어진 데이터를 표현한다. 마지막으로 학습자 메타데이터는 학습자의 초기 학습 관심정보와 학습한 경험에 의한 학습 콘텐츠에 대한 학습한 자료로 생성된 메타데이터를 토픽 맵으로 표현한다.

표 2. XTM을 이용한 Class 구축
Table 2. Class Building used to XTM

```

<topic id="C2_05">
  <baseName id="C2_06">
    <baseNameString>학습과정</baseNameString>
  </baseName>
</topic>
<topic id="C4_03">
  <baseName id="C4_04">
    <baseNameString>학습과목</baseNameString>
  ...
  <baseNameString>학습자</baseNameString>
  ...
    
```

표 2는 XTM을 이용한 Class 구축 샘플로, 어소시에이션 타입으로 학습과정, 학습과목, 학습자로 구성되며, 연결되는 어소시에이션 타입으로 직업, 나이, 학력, 선호강사, 선호과목, 선호분야, 수강비, 수강신청, 학습기간을 가지고 있다. 이를 위해 기존 클래스에 인스턴스를 삽입할 수 있도록 182개의 토픽과 15개의 연관관

계 타입 그리고 232개의 관계를 맷도록 구축하였다. 이는 그림 3과 같은 형태로 표현할 수 있다.

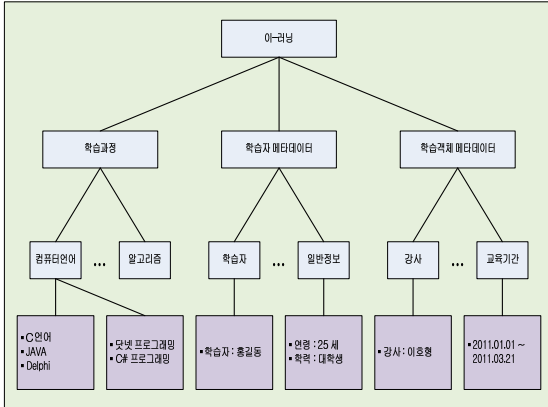


그림 3. 본 연구에서 온톨로지 구조
Fig. 3. Ontology Structure in this study

표 3은 XTM을 이용한 Association 구축 샘플로, 선호강사 및 수강비, 학습기간에 대한 연결을 표현하고 있다.

표 3. XTM을 이용한 Association 구축
Table 3. Association Building used to XTM

```

<topic id="T1_03">
  <baseName id="T1_04">
    <baseNameString>선호강사</baseNameString>
  </baseName>
</topic>
<topic id="M1_03">
  <baseName id="M1_04">
    <baseNameString>수강비</baseNameString>
  ...
  <baseNameString>학습기간</baseNameString>
</baseName>
</topic>
    
```

표 4는 XTM을 이용한 Topic 인스턴스 구축 샘플을 나타내고 있다.

표 4. XTM을 이용한 Topic 구축
Table 4. Topic Building used to XTM

```

<?xml version="1.0" encoding="UTF-8" ?>
...
- <instanceOf>
  <topicRef xlink:href="#X1_04" />
...
  <baseNameString>학습과목</baseNameString>
  </baseName>
...
  </instanceOf>
  <baseName id="S03_01">
    <baseNameString>Inside C#</baseNameString>
    </baseName>
...
  <baseName id="S03_03">
    <baseNameString>c#
    Programming</baseNameString>
    </baseName>
  </topic>
    
```

VI. 온톨로지 시각화 및 추천 시스템

구축된 온톨로지의 검색 및 추천을 위해 자바를 이용한 온톨로지 검색 및 시각화 시스템을 사용[13]하였으며, 구축된 토픽맵을 파서를 이용하여 트리플 형태로 추출하고 이를 활용하여 검색하였다. 여기서 82개의 토픽과 15개의 연관관계 타입 그리고 232개의 관계를 맷도록 구축된 정보를 토대로 총 415개의 트리플이 추출되었다.

표 5. 예상 시나리오
Table 5. Estimated Scenario

서울에 거주하는 B는 컴퓨터 프로그래밍에 관심이 많아 C++에 대한 과목을 마스터하고 다른 언어를 배우고자 한다. 기존에 학습했던 C++와 연계되는 언어를 배우고 싶으며 강사 A의 강의를 좋아한다.

B의 기존 학습 과목, 성과도 등이 메타데이터로 기록되어 있고 선호하는 강사가 A로 등록되어 있는 상황에서 C → C++ → C#으로 교과목 트리가 구성되어 있는 상황으로 본 경우에 A가 강의하고 있는 C# 관련 교과목이 추천된다. 그림 4에서 왼쪽에는 학습자가 기존 수강 과목과 메타데이터를 바탕으로 추천 과목을 보여주고 있으며, 학습 콘텐츠를 시각화하여 학습자가 관련 정보를 쉽게 검색 할 수 있도록 해준다.

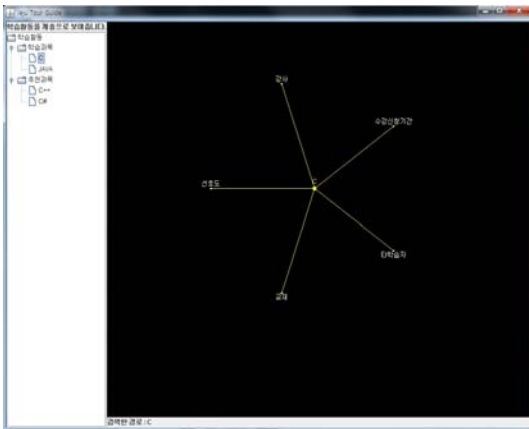


그림 4. 학습 콘텐츠의 검색 및 추천
Fig. 4. Retrieval and Recommendation of Learning Contents

여기서 알 수 있듯이, 구축된 온톨로지에 관해 정보가 풍부해야 하며, 추천 규칙에 따라서도 결과는 상

이하게 달라지기 때문에, 용도에 맞는 클래스와 인스턴스의 선택 및 적절한 추론 규칙의 선택이 중요하게 된다

VII. 결론 및 향후 연구

본 논문에서는 토픽 맵을 이용한 지능형 학습 콘텐츠 구축 방안에 관한 논문으로, 사용자 선호도를 이용한 메타데이터 구성 방안과 학습 콘텐츠의 토픽 맵 구축 방안에 대해 기술하였다. 이를 통해 6장과 같이 향후 사용자 선호도에 따른 콘텐츠 추천 시스템에 관해 연구할 계획이며 학습 객체들의 정보로 상호 연결되어 설계된 내비게이터를 기반으로 웹 인터페이스를 제공하여 학습자들이 보다 더 상세한 학습단위의 정보를 탐색하게 해주는 어플리케이션 프레임워크 (Application Framework)의 구현이 필요하다.

참고문헌

- [1] 남명준, "토픽 맵을 이용한 시소러스의 구조화 연구," 정보관리학회지, 제 22권 제 3호, 2005.
- [2] http://en.wikipedia.org/wiki/Topic_map
- [3] <http://isotopicmap.org/>
- [4] <http://www.topicmap.com/>
- [5] <http://www.topicmaps.org/xtm/index.html>
- [6] <http://www.topicmaps.org/xtm/index.html#desc-topic>
- [7] 하은옥, "토픽 맵 시각화 도구의 효율적 구조에 관한 연구," 상명대학교 대학원 석사 논문, 2002.
- [8] 김주섭, "토픽맵(Topic Map)을 이용한 e-Learning 시스템 온톨로지 개발에 관한 연구," 석사학위논문, 전북대학교, 2006.
- [9] 정경희, "토픽맵 기반 개인별 학습 콘텐츠 탐색 내비게이터 구조 설계," 석사학위논문, 조선대학교, 2008.
- [10] Myongho Yi, "Ontology Implementation and Methodology Revisited Using Topic Maps based Medical Information

- Retrieval System,” 한국정보관리학회 정보관리학회지, 제27권 제3호, pp. 35~51, 2010.
- [11] Strickert, M., Schleif, F.-M., Seiffert, U., Villmann, T., “Derivatives of Pearson correlation for gradient-based analysis of biomedical data”, *Inteligencia Artificial, Revista Iberoamericana de IA* 12(37), pp. 37-44, 2008
- [12] 오삼균, 박옥남, “토픽맵-기반 판소리 검색시스템 구축 및 평가에 대한 연구”, 한국도서관정보학회지, 제36권, 4호, pp. 77-98, 2005
- [13] Chang Choi et al. “Travel Ontology for Intelligent Recommendation System”, *Third Asia International Conference on Modeling & Simulation*, pp. 637-642, 2009,
- [14] AeLTC, "ADL SCORM Version 1.3 Application Profile", 2003.

저자소개



이호영(Ho-Young Lee)

1995년 조선대학교 대학원 컴퓨터공학과(공학석사)

2000년 조선대학교 대학원 컴퓨터공학과(공학박사)

2000년~현재 초당대학교 정보통신공학과

※ 관심분야 : 콘텐츠 구축, 네트워크, 시스템소프트웨어 등