

텍스트 내용 기반의 철학 온톨로지 구축 및 교육에의 응용

정현숙*, 최병일**

가톨릭대학교 컴퓨터정보공학부*, 서울대학교 철학사상연구소**

요 약

철학을 비롯한 인문학의 연구에 있어서 연구자들에게 텍스트는 지식을 습득하는 중요한 매체이다. 그러나 현재의 서지정보 중심의 문헌 검색으로는 연구자들이 여러 텍스트들로부터 필요한 지식만을 선택적으로 습득하는데 많은 시간과 노력을 들인다. 본 논문에서는 텍스트 내용에 의거한 지식 검색이 가능하도록 철학 고전 텍스트를 대상으로 철학 온톨로지를 구축하였다. 철학 온톨로지는 전문가에 의해 동서양의 철학 고전 텍스트로부터 분석된 주요 철학 개념과 의미 기반 계층관계 및 연관관계 정의하고 있으며 개념 해설서를 비롯한 온오프라인의 철학 지식 자원들을 연결하고 있다. 본 논문에서는 철학 온톨로지 구축 과정을 3개의 단계와 14개의 세부과정으로 나열함으로써 다른 분야의 학문 온톨로지 구축을 위한 하나의 가이드라인을 제시하고 있으며 철학 온톨로지의 응용 사례로서 대학의 철학 수업에서의 교수자와 학습자 사이의 철학 온톨로지에 기반한 지식 습득 및 교류가 가능함을 보였다.

Building a Philosophy Ontology based on Content of Texts and its Application to Learning

Hyun-Sook Chung*, Byung-II Choi**

The Catholic University, Department of Computer Science*

Seoul National University, Institute of Philosophy**

ABSTRACT

Researchers of humane studies including philosophy acquire knowledge from understanding of their texts. They spent a lot time and efforts to retrieve, read and understand many texts relevant to their research fields using a metadata-based text retrieval system. In this paper, we develop a philosophy ontology that enables researchers to retrieve knowledge in the content of texts of philosophy. Our philosophy ontology includes concepts and their hierarchical and associative relationships defined by philosophy researchers. We propose a methodology for constructing text-based ontology comprised of three phases and fourteen steps. This methodology may be used to construct another ontologies for learning. Also, we introduce a case study for applying our philosophy ontology to acquire and interchange knowledge of philosophy between a professor and students during philosophy classes.

1. 서론

정보기술의 발달과 웹의 등장으로 인해 인문학의 연구에 있어서도 기존의 종이 위주의 정적인 방식에서 인터넷에 저장된 디지털 콘텐츠를 활용하는 동적인 방식으로 변화하고 있다. 인문학 연구의 경우 과거의 연구 성과를 여러 각도로 조합함으로써 새롭게 해석하고 새로운 사실들을 발견해 나가는 과정으로 볼 수 있다. 이러한 측면에서 인문학 연구에 정보기술을 활용하는 것은 매우 바람직한 일이며 문학, 역사, 철학 등 다양한 분야에서 다양한 형태로 디지털 지식 자원을 구축하고 있다[1].

인문학의 한 분야인 철학에서는 철학 고전 텍스트들을 중심으로 학문 연구와 교육이 이루어지고 있다. 이에 따라 철학 연구자들은 자신의 연구와 관련된 지식을 얻기 위해 많은 철학 텍스트들을 읽고 이해하고 분석하는데 대부분의 시간과 노력을 들이고 있다. 이러한 텍스트 기반의 지식 습득과 지식 교류를 위한 현재의 정보 검색은 다음과 같은 문제점을 가지고 있다[2].

첫째, 철학 연구자들은 철학 문헌을 검색하기 위해 검색 엔진, 전자 도서관, 학술문헌 정보 포털 등을 이용함으로써 과거에 비해 짧은 시간에 관련 문헌 및 정보를 찾을 수 있다. 그러나 검색의 결과가 대부분 서지 정보만을 제공하기 때문에 텍스트의 내용 지식 습득은 여전히 어려운 문제이다.

둘째, 텍스트의 내용 지식 표현의 문제이다. 전자 도서관이나 학술 데이터베이스가 제공하는 대부분의 문헌 제공 형태는 HTML 또는 PDF 문서인데, 이는 매체만 달려졌을 뿐 지식 표현에 있어서 책과 동일하여 연구자들이 원하는 내용만 부분적으로 발췌하고자 할 경우에도 원문을 출력해서 전체를 읽어야 하는 문제를 가진다.

셋째, 검색된 문헌의 주요 개념들과 각 개념의 서술 구조를 파악하기 어렵다. 문헌의 목차가 그 문헌의 서술 구조를 보여주고 있으나 저자가 얘기하고자 하는 핵심 주제 및 관련 주제에 대한 구체적인 정보가 없으며 문단 수준의 상세한 문서 구조를 보기가 어렵다.

넷째, 검색된 여러 문헌들 사이에 있어서 내용의 동일, 대립, 보완 등 여러 가지 연관성에 대해 각각의 문헌을 읽고 비교하기 전에는 파악할 수 없다. 문헌의 제목과 간단한 요약문 정도로는 연구자가 원하는 문헌을 결정하기 어렵다.

마지막으로 문헌내의 특정 용어 또는 개념에 대해 이와 관련된 인터넷 상의 유용한 지식 자원들이 연결되어 있지 않으므로 지식 검색이 제한적이다.

이러한 문제점들의 근본적인 원인은 현재의 문헌 정

보화 서비스가 도서관 서고에 보관되는 출력물을 단순히 컴퓨터에 저장 가능한 전자문서로 변환하여 데이터베이스화해 놓은 것에 불과하기 때문이다. 그러므로 이에 대한 해결책은 분야별 전문가에 의해 텍스트가 가지는 주요 개념들과 개념을 설명하는 서술 구조를 분해하여 이 개념 요소를 기반으로 지식을 조직하고 검색 가능하도록 하는 것이다.

본 논문에서는 분야별 철학 전문가들에 의해 동양과 서양의 주요 철학 고전 텍스트를 분석하고 해설한 내용을 기반으로 철학 지식을 구성하는 철학 온톨로지(philosophy ontology)를 설계 및 구현하였다.

철학 온톨로지의 확장성을 고려하여 3계층의 온톨로지 구조를 설계하였으며 이전 온톨로지 구축 방법론을 확장하여 철학 텍스트 분석에 의한 온톨로지 구축 절차를 제안하고 있다. 온톨로지 구조 및 구축 절차는 온톨로지 비전문가인 철학 전문가들에게 온톨로지 구축에 적합하도록 텍스트 분석을 유도하는 하나의 지침(guideline)으로 사용된다.

또한 본 논문에서는 철학 온톨로지를 이용한 지식 습득 및 교류의 응용 분야로서 대학의 철학 교육에의 적용 가능성을 보이고 있다. 대학에서의 철학 교육은 대부분 교수가 제시하는 주요 개념들을 학생들이 문헌을 조사하여 파악된 지식을 발표하고 이에 대하여 교수가 오류를 지적하며 포괄적인 개념 설명과 함께 관련된 다른 저술의 내용을 연관 지어 설명하는 방식이다. 이러한 텍스트 내용 지식 기반의 교수-학습 상황에서 철학 온톨로지는 지식 검색, 습득, 교류의 길잡이 역할을 한다.

본 논문의 구조는 다음과 같다. 2장에서는 본 연구와 관련있는 이전 연구들을 살펴보고 3장에서는 철학 온톨로지의 계층적 구조를 설명한다. 4장에서는 철학 온톨로지 구축 절차를 과정별로 소개하고 5장에서는 철학 온톨로지를 이용한 철학 교육의 가능성에 대해 설명한다. 6장에서는 철학 온톨로지의 의의와 향후 연구 방향을 소개하며 결론을 맺는다.

2. 관련연구

관련연구로서 먼저 철학 연구자들이 자주 사용하는 철학 디지털 지식 자원의 형태를 살펴보면 크게 다음의 4가지 유형으로 분류할 수 있다[4]. 철학 디지털 텍스트 사이트[12]는 주로 철학 텍스트들의 원전 또는 번역본을 직접 또는 하이퍼링크를 통해 제공하는 웹사이트들이다. 이와 같은 웹사이트들은 철학 연구자들이 1차 문헌을 웹을 통해 직접 읽거나 인용할 수 있도록 하기 위한 것이다.

철학 디지털 저널 사이트[10]는 주로 철학 논문들을 직접 또는 하이퍼링크를 통해 제공하는 웹사이트들이다. 이와 같은 웹사이트들은 철학자들의 연구 활동의 핵심이 되는 논문 출판과 논문 인용을 웹에서 이루어질 수 있도록 지원하기 위한 것이다.

철학 디지털 참고문헌 사이트[15]는 주로 철학 디지털 사전 또는 철학 디지털 백과사전을 제공하는 웹사이트들이다. 이와 같은 웹사이트들은 철학 연구자들이 주로 사용하거나 참고하는 용어 및 철학자, 철학 문헌, 철학 개념 등에 대한 지식을 웹에서 찾을 수 있도록 하기 위한 것이다.

철학 메타 사이트[13]들은 철학 관련 사이트들에 대한 정보 및 디렉토리를 제공하는 웹사이트들이다. 이와 같은 웹사이트들은 철학 텍스트, 철학 저널, 철학 단체, 철학자 개인 뿐 아니라 철학 토픽, 철학 분야 등에 대한 철학 지식을 제공하는 철학 웹사이트들에 대한 메타 정보를 제공하기 위한 것이다.

이들 철학 지식 제공 사이트들과 달리 철학 온톨로지는 원문 제공을 목적으로 하는 것이 아니라 원문의 내용을 분석하여 중요한 개념들을 추출하고 각 개념들에 적합한 용어를 정의하며 이들 용어사이에 연관 관계를 정의하는 의미적 지식망을 구축하는 것이다.

온톨로지를 효율적으로 구축하기 위한 연구로서 온톨로지 공학적 측면의 방법론이 연구되어 왔으며 그 접근법은 크게 순공학(Forward engineering)과 역공학(Reverse engineering)으로 나누어진다.

순공학적 접근법은 전문가에 의한 지식 분석 및 추출로부터 온톨로지를 구축하고 온톨로지와 지식 자원을 연결하는 절차로서 대표적 온톨로지 구축 방법론으로 엔터프라이즈 온톨로지 방법론(Enterprise Methodology)[17], TOVE(TORonto Virtual Enterprise) 방법론[6], METHONTOLOGY[7] 방법론, On-To-Knowledge[14] 및 Development 101[9]이 있다. 이들 방법론은 대부분 기업, 화학공정 등 특정 도메인에서의 온톨로지 구축 경험을 기반으로 작성된 방법론으로서 상위 수준의 관점에서 온톨로지 생성 절차를 단계별로 정의하고 있다[8].

역공학적 접근법[10][11]은 구조적, 반구조적 및 자유 형식의 지식자원에서부터 자연어처리에 의하여 명사형 및 동사형 용어들을 추출한 다음 기계 학습 알고리즘으로 온톨로지를 구축하는 반자동적인 절차이다. 그러나 역공학적 접근법을 적용할 수 있는 도메인은 지식자원의 형식이 온톨로지로 변환하기에 적합한 구조를 가지거나 아니면 상품정보, 용어사전 등과 같이 명사형 용어들을 쉽게 추출할 수 있으며 의미적으로 명확한 경우이다.

철학 온톨로지 구축 절차는 순공학적 접근법을 따르며 텍스트 내용 분석에 의한 개념 추출에 적합하도록 개념화 및 구현 단계의 세부절차를 정의하였다.

3. 철학 온톨로지 구조

철학 온톨로지의 전체적인 구조를 보면 참조 온톨로지, 철학 분류 온톨로지, 철학 텍스트 온톨로지의 3 계층으로 구성되어 있다. 참조 온톨로지는 스키마 온톨로지로서 온톨로지를 구성하는 요소의 주체성(identity) 정의, 속성 정의, 연관성 정의 및 제약조건 등을 정의하고 있으며 철학 분야별 소규모 온톨로지들이 공통적으로 참조하는 상위 수준의 온톨로지이다. 철학 분류 온톨로지는 철학 분야의 일반적인 분류에 따른 개념 정의와 개념간 계층관계를 정의하고 있는 온톨로지이다. 철학 일반적인 분류는 대부분의 철학 사전, 철학 웹사이트 등에서 제공하는 분류로서, 철학자, 철학문헌, 철학이론, 철학용어, 철학분야, 철학학과 등이다. 철학 온톨로지에서는 동양과 서양의 지역적 분류와 고대, 중세, 근대, 현대의 시대적 분류, 그리고 철학자, 철학문헌 등의 철학 요소 분류에 따라 계층적으로 철학을 분류하고 있다.

철학 텍스트 온톨로지는 각 철학 고전 텍스트별로 하나의 단위 온톨로지를 생성한 다음 이들을 지역적, 시대적 및 철학 요소별로 분류한 온톨로지이다. 텍스트 단위 온톨로지는 해당 텍스트의 최상위 주요 개념들을 중심으로 그 하위 개념들을 내포-외연 관계에 따라 계층적으로 구성하고 개념들 사이에 동일관계, 대립관계, 보완관계 등의 의미적 연관관계를 설정하고 있다. 또한 각 개념들의 자원으로 고전 텍스트의 원문의 일부분 및 해설서의 일부분을 링크하고 있다.

그림 1은 철학 온톨로지의 3 계층 구조를 보이고 있다. 지식층이 철학 온톨로지이고 자원층은 고전 텍스트 및 해설서의 HTML 파일, PDF 파일, 철학 관련 이미지, 동영상, 다른 웹사이트의 페이지 등을 가지는 철학 지식 자원들이다.

그림 1에서 임마누엘 칸트는 서양근대철학자와 IS-A 관계를 가지며 속성으로 생애해설을 가진다. 실천이성비판은 서양근대철학문헌과 IS-A 관계를 가지며 텍스트 내용의 주 개념으로 이성, 도덕법칙, 준칙 등을 가진다. 이성은 다시 이성의 정의, 이성의 구분과 PART-OF 관계를 가지며 이성의 구분은 속성으로 해설과 인용을 가진다. 그리고 이 해설과 인용 속성 값은 고전 텍스트 및 해설서 HTML 파일의 일부 영역과 연결된 URI 주소 값이다.

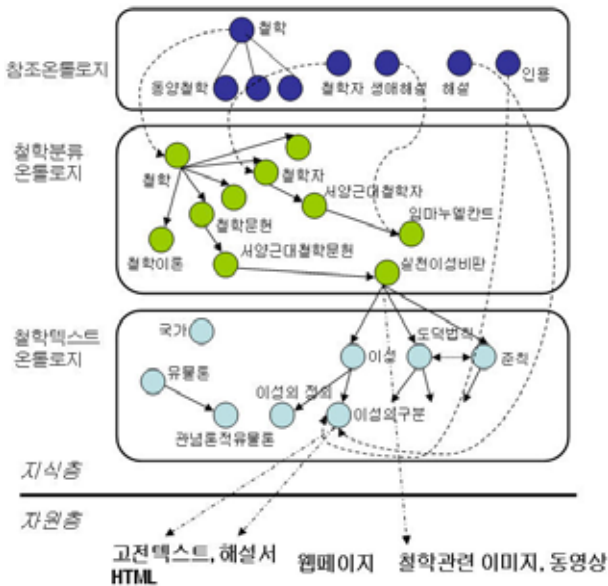


그림 1. 철학 온톨로지의 3 계층 구조

4. 철학 온톨로지 구축 절차

본 장에서는 철학 온톨로지의 생성 과정을 소개하고 각 과정에서의 중요한 고려 사항들에 대해 서술한다. 특히 온톨로지 구축 측면에서 철학 텍스트를 분석하고 개념을 추출하는 방법에 대해 자세하게 정의한다.

철학 온톨로지 생성 과정은 그림 2에서 볼 수 있듯이 크게 준비 단계, 개념화 단계, 구현 단계로 나누어지고 개념화 단계에서는 철학 포괄적 지식 정의 과정과 철학 텍스트의 지식 정의 과정으로 분류된다. 그리고 구현 단계에서는 토픽맵(Topic Maps)[3,5]을 기반으로 철학 온톨로지를 구현하는 과정들로 구성된다.

4.1 준비 단계

준비 단계는 철학 온톨로지 구축 목적 및 범위 설정 과정과 전문가팀 구성 및 자료 수집 과정으로 구성된다. 철학 온톨로지의 범위는 동서양 및 고대에서 현대까지의 철학 전반에 걸쳐 중요 고전 텍스트들을 대상으로 하고 있으며 약 300여권의 대상 텍스트를 선정하고 있다.

철학 온톨로지의 전문가팀은 크게 문헌 분석 팀과 온톨로지 구현 팀으로 나누어져 있으며 문헌 분석 팀은 철학의 각 분야별 전문 지식을 보유하고 있는 전문가들로 구성되어 있다. 이들은 주로 문헌에서 주요 개념들을 추출하고 표준 용어를 정의하며 개념들 사이의 관계를 설정한다. 또한 고전을 보다 쉽게 이해할 수 있도록 각 개념들에 대해 객관적인 입장에서 해설과

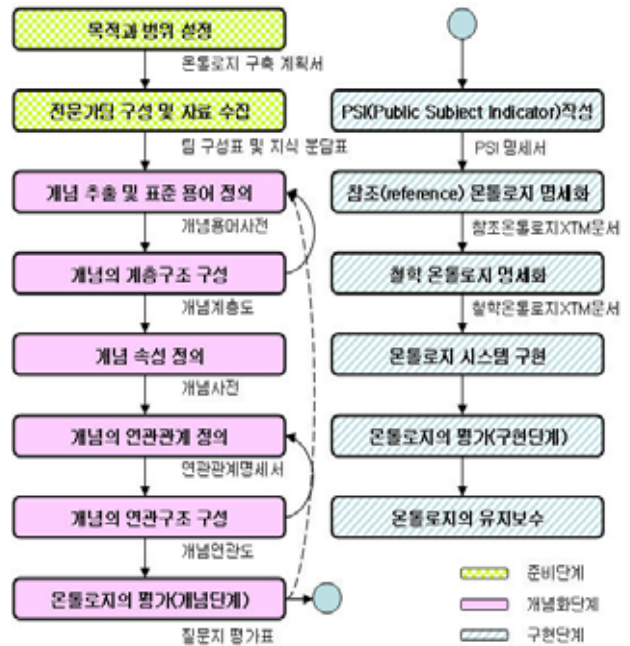


그림 2. 철학 온톨로지 구축 절차

인용을 기술한다. 온톨로지 구현 팀은 전문가가 분석한 개념들을 토픽맵으로 구현하는 팀으로서 컴퓨터상에 온톨로지를 구현하고 테스트하며 문헌외의 유용한 지식 자원들을 연결함으로써 온톨로지를 완성하는 일을 담당한다.

4.2 개념화 단계

개념화 단계는 문헌 분석, 개념 추출, 용어 정의, 관계 정의 등 철학 온톨로지를 비형식적인 언어로 생성하는 단계이다. 비형식적인 언어는 그림, 문자 등의 비구조화된 형식을 말한다. 철학 온톨로지의 개념화 단계는 철학 전반의 포괄적 지식의 개념화와 철학 텍스트 내용 지식의 개념화로 나누어진다.

4.2.1 철학의 포괄적 지식 개념화

온톨로지를 생성하기 위해 특정 도메인의 지식을 분석하고 추출하는 기법으로 하향식 기법(top-down approach), 상향식 기법(bottom-up approach), 그리고 상하향식 기법(middle-out approach)[8]이 있다. 이 기법들 중에서 최상위 수준의 개념으로부터 출발하여 상세 개념들로 분류해 나가는 기법이 하향식 기법이며 철학 온톨로지에서도 철학 전반의 포괄적 지식 구조를 생성하기 위해 사용된다.

철학 온톨로지에서도 최상위 개념은 철학(philosophy)이다. 그리고 그 하위에 철학 개념과 'PART-OF' 관계

를 가지는 하위 개념으로 철학자, 철학문헌, 철학분야, 철학사, 철학학과, 철학이론, 철학용어의 7가지 세부 개념들을 분류하였다. 그리고 철학자의 경우 지역적 및 시대적 기준에 따라 철학자와 'is-a' 관계를 가지는 한국철학자, 중국철학자, 인도철학자, 서양고대철학자, 서양중세철학자, 서양근대철학자, 서양현대철학자로 분류하였으며 철학문헌, 철학사 등 다른 세부 개념들도 이와 동일한 기준으로 분류하였다.

온톨로지에서 개념들 사이의 관계는 위에서 언급한 'IS-A'와 'PART-OF'의 수직적 계층 관계 외에 의미정보를 기반으로 연결되는 수평적 연관 관계로 나누어진다. 수평적 연관 관계는 개념들 사이의 의미적 연관성을 표현하는 관계로서 철학 전문가가 가지는 지식에 의해 구체화된다. 예를 들어, 철학자 개념과 철학문헌 개념 사이의 연관 관계는 '저자' 또는 '주요 저서', '대표 저서' 등이 존재하고 철학자와 철학이론과의 연관 관계는 '주장이론', '반대이론' 등의 관계가 존재한다. 철학자 사이에는 '스승과 제자', '영향을 받은 철학자', '영향을 끼친 철학자' 등의 관계가 성립될 수 있다.

철학 온톨로지 생성 프로세스의 개념 분석 및 추출 과정의 산출물은 개념사전(concept dictionary)이다. 개념사전은 철학 온톨로지의 개념들의 집합이며 하나의 개념은 그림 3과 같이 정의된다.

철학 온톨로지에서는 하나의 개념은 하나의 CID(concept identifier)를 가진다. CID는 유일한 값으로 개념들 사이의 참조에 사용된다. 레벨(level)은 계층 구조에서의 위치를 가리키는 것으로서 최상위 개념 철학의 레벨이 1이고 그 하위의 철학자는 철학의 하위 개념들 중의 첫 번째 개념이므로 레벨이 1.1로 표기된다. 개념 타입(concept type)은 철학 온톨로지에서는 개념의 성격을 가리키는 것으로 철학자는 그 하위에 다른 개념들을 포함하며 직접 지식 객체를 가리키는 것이 아니므로 클래스가 된다. 반면, 지식 객체에 대한 내적 또는 외적 참조가 있으면서 하위 개념을 가지지 않는 온톨로지에서는 최하위 개념들의 개념 유형은 인스턴스(instance)가 된다.

슈퍼 클래스는 철학자 개념을 내포하는 상위 개념의 CID를 가리키고 서브 클래스는 철학자가 내포하는 하위의 개념들의 CID를 가리킨다. 이때 슈퍼 클래스와 서브 클래스 값은 0개 이상 여러 개가 존재한다. 속성(attributes)은 철학자 개념을 설명하는 항목으로서 0개 이상 여러 개가 존재하며 속성명, 내적/외적 참조, 다중성, 속성 타입을 가진다.

관계성(relationships)은 철학자 개념과 다른 개념들과의 사이에 존재하는 연관 관계에 대해 정의하는 것으로 관계명(name), 기준 개념(origin), 대상 개념

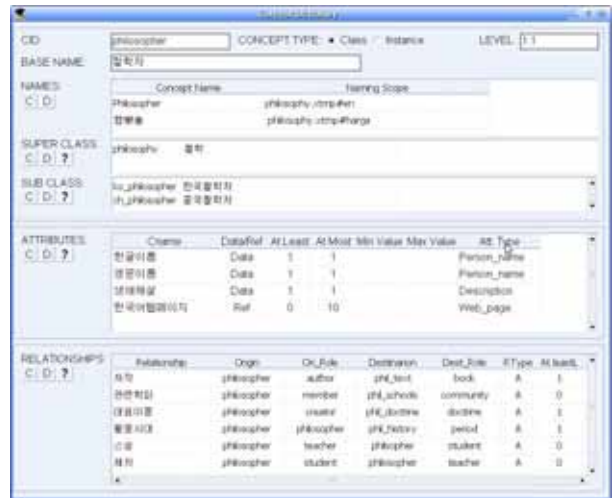


그림 3. 온톨로지 편집기의 개념 정의 화면

(destination), 기준 개념의 역할(ori_role), 대상 개념의 역할(dest_role)을 가진다. 이때 기준 개념과 대상 개념은 모두 CID로 표시되며 속성 타입과 역할도 하나의 개념으로 정의되므로 각각의 CID로 표시된다.

4.2.2 철학 텍스트의 내용 지식 개념화

철학 온톨로지의 상위 수준의 온톨로지는 철학 전반의 포괄적인 지식을 정의하고 분류하는 온톨로지인 반면 철학 텍스트 온톨로지는 텍스트 내의 내용상에 존재하는 지식을 정의하고 분류하는 내용 기반 온톨로지이다. 그림 4는 철학 온톨로지의 두 계층 사이의 연결을 보이고 있다.

내용상의 지식 추출을 위한 텍스트 내용 분석 기법은 상하향식 기법을 사용하였다. 텍스트 분석을 위해 하향식 기법을 사용할 경우의 문제점은 추출된 개념 집합 속에 텍스트 내용상의 개념보다 상위 수준의 일반적인 개념들이 더 많은 부분을 차지할 수 있다는 것이다.



그림 4. 철학 온톨로지의 계층간 연결구조

그리고 상향식 기법을 사용할 경우의 문제점은 최하위 인스턴스 타입의 개념에서 상위의 클래스 타입의 개념들을 추출하여 올라갈 경우 상위의 개념들이 너무 넓게 파악될 수 있다는 문제점을 가진다. 이에 반해 상하향식 기법은 텍스트의 주요 개념에서 출발한다. 모든 철학 텍스트는 저자가 강조하고 설명하는 주요 개념들이 존재하므로 이 개념들을 시작점으로 두고 아래로는 하향식 기법을 적용하고 위로는 상향식 기법을 적용하여 계층 관계의 개념들을 추출한다. 그림 5는 이 세 가지 기법의 처리 형식을 보여준다.

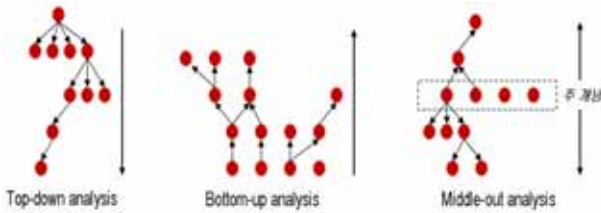


그림 5. 철학 텍스트의 내용상의 지식 분석 및 추출 기법

텍스트에서 추출되는 개념의 이름은 단일명사, 복합명사, 명사구 형태를 가진다. 이 경우 개념명이 명사구 형식일 경우에는 개념명을 아래 규칙에 따라 분석함으로써 새로운 개념 및 계층 구조를 파악할 수 있다. 본 논문에서는 대표적인 명사구 유형 4가지를 제시하고 각각의 경우에 어떤 규칙으로 분석이 가능한지 보인다.

① 규칙 1. 개념명 형식이 'A적 B'인 경우

규칙 1의 예시는 '법적 관점', '자연주의적 오류', '도덕적 소외', '관념론적 역사관' 등으로 볼 수 있으며 이 경우 B가 문서내의 중요 개념일 경우 반드시 B를 상위 개념으로 분류하고 그 하위에 'A적 B' 개념을 두어야 한다. 예를 들어, '역사관' 개념을 별도로 분류하고 그 하위에 '관념론적 역사관'을 연결하여야 한다.

그러나 '자연주의적 오류'와 같이 오류 자체가 문서 내에서 중요하게 서술되는 개념이 아닐 경우 오류를 별도로 분류하지 않고 '자연주의' 하위에 '자연주의적 오류'로만 분류한다.

② 규칙 2. 개념명 형식이 'A의 B(B of A)'인 경우

규칙 2의 예시는 '물질의 증여', '도덕법칙의 원천', '행위의 자유', '의지의 자유' 등으로 볼 수 있으며 이 경우 A가 중요한 개념이라면 A를 상위 개념으로 분류하고 그 하위에 'A의 B'를 두며 반대로 B가 중요한 개념이라면 B를 상위 개념으로 분류하고 그 하위에 'A의 B'를 둔다. 만일 A와 B가 각각 중심 개념으로 내용 서술을 가진다면 A와 B를 상위 개념으로 두고 그 하위에 'A의 B'를 둔다. 이때 'A의 B'는 클래스의 다중상속처럼 여러 상위 개념을 가진다.

③ 규칙 3. 개념명 형식이 'A와(과) B', 'A와(과)B의 C', 'A의 B와(과) C'인 경우

규칙 3의 예시는 '기도와 책임', '선과 양심', '자유와 초탈의 복음', '의지의 주관성과 객관성' 등으로 볼 수 있으며 'A와(과) B'의 경우 A와 B를 각각 상위 개념으로 분류하고 그 하위에 A와 B의 관계에 대해서 설명하는 자세히 분류된 여러 개의 하위 개념들을 생성하여 계층 구조로 연결한다. 'A와(과) B의 C'의 경우에는 C를 상위 개념으로 분류하고 그 하위 개념으로 'A의 C'와 'B의 C'를 둔다. 예를 들어, '자유와 초탈의 복음'의 경우 '복음'이라는 넓은 개념 속에 '자유와 초탈의 복음'과 '초탈의 복음'이라는 좁은 개념들이 존재하는 것으로 정의할 수 있다. 또한 'A의 C'와 'B의 C' 사이에 의미적 관계 설정을 위해 연관 관계를 정의한다. 'A의 B와(과)C'의 경우에는 A를 상위 개념으로 분류하고 그 하위 개념으로 'A의 B'와 'A의 C'를 연결한다. 이때 'A의 B'와 'A의 C' 사이에 의미적 관계 설정을 위해 연관 관계를 정의한다.

④ 규칙 4. 개념명 형식이 'A(으)로서의 B'인 경우

규칙 4의 예시는 '법의 지반으로서의 자유의지', '자유로서의 법' 등으로 볼 수 있으며 이 경우 B를 상위 개념으로 분류하고 그 하위 개념으로 'A(으)로서의 B'를 연결한다.

이들 규칙에서의 원칙은 조사에 의해 제한이 되는 용어를 상위 개념으로 분류하고 하위 개념과의 사이에 계층 구조를 정의한다는 것이다. 또한 대등관계의 용어들 사이에는 어떠한 의미적 연관성이 있으므로 그 연관성에 대해 또 다른 하위 개념들을 정의하거나 또는 연관 관계를 정의하여야 한다.

철학 텍스트 온톨로지는 개념 사전 대신에 그림 6과 7에서 볼 수 있는 개념계층도와 개념연관도를 작성하였다. 하나의 텍스트에서 추출되는 개념의 수가 최소 300개 이상이므로 동일한 형식을 가지는 개념들에 대해 개념 사전을 개별적으로 작성하는 것은 비효율적이다.

그림 6의 개념계층도는 분석된 개념의 개념명과 CID, 계층 구조 그리고 개념의 철학 지식이 존재하는 텍스트의 위치를 내용으로 작성된다. 개념의 CID는 상위 수준의 개념과 달리 의미적인 이름을 부여하기 어렵기 때문에 레벨과 일련번호 중심으로 부여된다. 즉 CID 자체에 온톨로지의 계층 구조에서의 위치 정보를 가지도록 생성 규칙을 정하였다.

개념의 철학 지식의 존재 위치를 가리키는 resource ID는 문단일 경우 시작 문자를 p로 하였으며 인용문일 경우에는 시작 문자를 q로 하였다. p나 q 뒤의 일련번호는 텍스트의 구조를 가리킨다. 즉 1부, 1장, 1절, 1문

해결·법철학·개념 계층도		
CID	Concepts Hierarchy	Resource ID
c1	의지	
c11	의지의 세계기	PL1.1.2.1
c12	의지의 세형태	PL1.1.3.1
c121	자연적 의지	PL1.1.3.3
c122	자의	PL1.1.3.5
c123	숙자 대자적 의지	PL1.1.3.8
c13	의지의 주관성	PL1.1.4.1
c131	의지의 주관성의 의미	PL1.1.4.3
c14	의지의 객관성	PL1.1.5.1
c141	의지의 객관성의 의미	PL1.1.5.2
c2	자유	
c21	자유의지	
c211	법의 기반으로서의 자유의지	PL1.1.1.4
c212	자유로서의 법	PL1.1.1.9
c3	소유	
c31	소유의 계급성	
c311	점유취득	PL1.2.2.1.1
c3111	점유취득의 방법	PL1.2.2.1.4
c31112	목적적 취득	PL1.2.2.1.5
c31113	형식의 부여	PL1.2.2.1.7
c31114	표기	PL1.2.2.1.10

그림 6. 철학 텍스트의 개념 추출 과정의 산출물인 개념계층도

단인 경우 p1.1.1.1로 명명된다. 개념에 의해 참조되는 텍스트의 범위는 최소 문단 단위로 하였으며 인용문인 경우에는 문단내의 한 인용문을 최소 단위로 하였다.

그림 7의 개념연관도는 철학 텍스트의 개념들 사이의 의미적 연관성을 보이는 산출물로서 텍스트를 분석하는 철학전문가의 지식에 의존하여 분석된다. 텍스트 내에서 개념들 사이의 연관성은 그 자체가 하나의 중요한 사상이고 개념이 된다. 예를 들어, 칸트의 실천이성비판에서 이성과 도덕법칙의 관계는 한 단어로 설명되지 않는 그 자체가 여러 페이지에 나누어 설명되는 중요한 개념인 것이다. 이 경우 이성과 도덕법칙 사이에는 일반적 연관성을 규정하기 어려우므로 연관 관계로 표현하지 않고 두 개념 사이의 관계를 설명하는 여러 하위 개념들을 추출한 다음 개념들 사이에 계층 구조를 정의해야 한다.

4.2.3 구현 단계

철학전문가에 의해 생성된 철학 온톨로지를 토픽맵으로 구현하기 위해 먼저 철학 PSI(Published Subject Indicator)와 철학 스키마를 정의하였다. 철학 PSI는 철학 온톨로지의 주요 용어에 대한 정의를 포함하는 것으로 철학, 철학자, 철학문헌, 철학이론 등의 개념에 주체성(identity)을 부여한다. PSI는 여러 개의 토픽맵들을 하나의 토픽맵으로 통합할 때 동일한 주체성에 속하는 토픽들을 그룹화하기 위해 사용된다.

철학 스키마는 클래스 토픽을 정의하고 토픽들 사이의 계층 구조를 정의하는 것으로 모든 인스턴스 토픽들이 공유할 수 있도록 토픽 타입, 어커런스 타입, 연관 관계 타입을 정의한다. 또한 철학 스키마는 인스턴스 토픽맵의 템플릿을 제공한다. 즉, 철학자 템플릿은 칸트, 헤겔, 데카르트 등 모든 철학자가 공유할 수 있는 템플릿이고 철학문헌 템플릿은 실천이성비판, 법철

학, 독일이데올로기 등 모든 철학 문헌들이 공유할 수 있는 템플릿이다.

개념화 단계에서 완성된 비형식 언어의 온톨로지를 컴퓨터상에 구현하기 위해서는 컴퓨터가 이해할 수 있는 형식 언어로 재 작성하여야 한다. 토픽맵의 형식 언어는 XML 형식의 XTM(XML Topic Maps)[16]이다. 그림 8은 XTM 구문으로 작성된 헤겔 철학자 토픽맵과 칸트 철학자 토픽맵의 일부분을 보이고 있다. 각각 헤겔과 칸트 토픽을 정의하고 있으며 헤겔과 칸트 사이에 'teacher_of'의 연관 관계가 존재함을 <association> 엘리먼트에서 정의하고 있다.

```

Hegel.xtm
<topic id="Hegel">
  <instanceOf <topicRef id="http://philosophy.dsp/inst_philosopher"/> <instanceOf>
    <setName> <setNameString>헤겔</setNameString> </setName>
    <occurrence>
      <instanceOf <topicRef id="http://philosophy.dsp/inst_philosopher"/> <instanceOf>
        <resourceData> 헤겔은 1788년 8월 27일경 신학 대학에 신학과 학생으로 입학했다. ... </resourceData>
        <occurrence>
          ...
        </occurrence>
      </instanceOf>
    </instanceOf>
  </topic>
  <association>
    <instanceOf <topicRef id="http://philosophy.dsp/teacher_of"/> <instanceOf>
      <member>
        <instanceOf <topicRef id="http://philosophy.dsp/teacher_of"/> </instanceOf>
        <instanceOf <topicRef id="http://philosophy.dsp/teacher_of"/> </instanceOf>
      </member>
    </instanceOf>
  </association>
</topic>

Kant.xtm
<topic id="Kant">
  <instanceOf <topicRef id="http://philosophy.dsp/inst_philosopher"/> <instanceOf>
    <setName> <setNameString>칸트</setNameString> </setName>
    </instanceOf>
  </topic>
</topic>

```

그림 8. 헤겔 철학자와 칸트 철학자의 정의 및 헤겔과 칸트 사이의 연관 관계의 정의

이와 같은 XTM 구문의 토픽맵 기술 문서를 컴퓨터상에 온톨로지로 구현하기 위해서는 XTM 문서를 읽고 해석한 다음 메모리내에 그래프 구조의 온톨로지를 구현할 수 있는 토픽맵 응용 시스템이 필요하다. 본 논문에서는 그림 9와 같은 구조를 가지는 시스템을 구현하였다.

시스템의 주요 구성요소는 사용자 인터페이스, 토픽맵 생성 및 검색 엔진, 토픽맵 저장소, 토픽맵 API 등의 4가지로 나누어볼 수 있다. 사용자 인터페이스는 JSP와 스트럿츠(struts)로 구현된 웹 인터페이스를 제공하는 것으로 토픽맵의 요소를 추가, 변경, 삭제할 수 있는 편집기(Topic Map Editor)와 토픽맵의 내용을 텍스트 기반으로 검색하는 토픽맵 네비게이터(Topic Map Navigator) 및 그래픽 기반으로 검색하는 비주얼 네비게이터(Visual Navigator)로 구성되어 있다.

토픽맵 엔진은 토픽맵 생성을 위한 부분과 생성된 토픽맵의 내용 검색을 위한 부분으로 나누어지는데, 생성은 토픽맵 파서(Topic Map parser), 정당성 검사기(Topic Map Validator), 구축기(Topic Map Factory)로 구성되고 검색은 색인의 생성 및 관리를 담당하는 색인 관리자(Index Manager), 키워드 검색을 지원하는

많지 않아서 그 내용을 이해하기가 쉽지 않다는 것이다. 또한 자신의 생각을 글로 작성한 다음 처음부터 읽어 내려가는 발표 방식이 의견을 주고 받는 토론에는 부적절하다는 것이다.

표 2. 철학 온톨로지 적용에 관한 설문내용
 ①매우그렇다②그렇다③보통이다④그렇지않다⑤매우그렇지않다

번호	설문항목	선택항목				
		①	②	③	④	⑤
1	과제의 주제선정이적절한가	①	②	③	④	⑤
2	자료조사를 위해 도서관을 이용하는가	①	②	③	④	⑤
3	자료조사를 위해 인터넷을 이용하는가	①	②	③	④	⑤
4	주제에 적합한 자료 수집이 용이한가	①	②	③	④	⑤
5	철학문헌의 해설이 풍부한가	①	②	③	④	⑤
6	글쓰기만으로 자신의 생각을 표현하기에 충분한가	①	②	③	④	⑤
7	자신의 글을 읽는 방식의 발표가 토론에 적절한가	①	②	③	④	⑤
8	철학온톨로지의 구조가 이해하기 쉬운가	①	②	③	④	⑤
9	철학온톨로지의 검색방식이 적절한가	①	②	③	④	⑤
10	주제와 관련된 검색된 자료가 충분한가	①	②	③	④	⑤
11	검색된 자료의 해설이 풍부한가	①	②	③	④	⑤
12	자료들 사이에 계층적 및 의미적 연관성이 풍부한가	①	②	③	④	⑤
13	자료들 사이의 관계가 개념 이해에 도움이 되는가	①	②	③	④	⑤
14	자신의 생각을 온톨로지로 표현하기가 용이한가	①	②	③	④	⑤
15	온톨로지에 근거한 발표 및 토론이 적절한가	①	②	③	④	⑤

이에 반해 철학 온톨로지를 활용한 방식에서는 먼저 주제와 관련된 자료를 검색하기가 용이하고 자료들 사이에 계층적 및 의미적 연관성에 따라 연속적으로 관련 자료들을 검색할 수 있다는 데에서 장점을 찾을 수 있다. 발표 또한 온톨로지를 보이고 상위 개념에서부터 세부 개념으로 내용 전개하듯이 발표함으로써 발표자와 다른 학생들 사이에 토론을 유도할 수 있는 시각적인 도구로서 활용된다. 단지, 학생들이 자신의 생각을 온톨로지 표현하는데 익숙하지 않기 때문에 글쓰기외에 많은 시간을 온톨로지 표현에 할애하는 어려움이 있다. 이러한 설문 결과를 볼 때 철학 온톨로지의 교육적 적용이 기존의 문제점을 해결하는 하나의 방법이 될 수 있음을 알 수 있었다.

이와 함께 학습자에 대한 평가 또한 객관적이며 교수자와 학습자들 사이에 공통된 견해를 가질 수 있다.

그리고 교수자는 학습자가 작성한 온톨로지의 오류를 수업 과정 중에 수정하면서 올바른 주장을 이끌어 낼 수 있다. 이러한 과정에 의해 생성된 온톨로지는 지속적으로 보관함으로써 향후 진행되는 철학 수업의 중요한 학습 자료로서 뿐만 아니라 철학 분야의 지식 저장소로 활용될 수 있다. 또한 학습자들에게는 자신의 지식이 저장되고 재사용됨을 인식시킴으로써 보다 충실한 조사와 발표가 이루어지도록 동기 부여할 수 있다.

7. 결론 및 향후연구

본 논문에서는 철학 텍스트에 들어 있는 지식을 개념 단위로 추출하고 표준 용어로 정의한 다음 이들 사이에 의미적 연관관계를 설정함으로써 철학 온톨로지를 구축하였다. 본 논문의 의의는 철학이라는 학문도메인의 실제적인 온톨로지를 구축하였다는 것과 철학 텍스트의 메타데이터뿐만 아니라 텍스트 내용을 분석하여 개념화하였다는 데에 있다. 또한 온톨로지 구현 과정을 세분화하여 소개하였으며 이러한 온톨로지 구현 과정은 여러 다른 학문분야에도 동일하게 적용될 수 있다.

현재의 철학 온톨로지는 동서양의 철학 고전 60여권을 대상으로 구축되어 있으며 향후 300여 권까지 확대할 예정이다. 그리고 철학 백과사전등과의 연계를 통하여 철학 전반의 지식을 체계적으로 보이고자 한다.

본 논문에서는 철학 온톨로지의 응용 사례로서 철학 수업에서의 지식 습득 및 지식 교류에의 활용을 보이고 있다. 철학 수업은 전통적으로 철학 개념을 소개하고 이에 대한 반론 또는 검증은 철학사별, 철학자별, 사상별로 전개해 나간다. 이러한 과정 속에 학습자는 주어진 토론 주제에 대해 자료를 조사하고 정리하여 자신의 지식을 글로 작성한 다음 이를 수업시간에 발표하고 토론한다.

이러한 철학 수업에서 철학 온톨로지는 학습자에게 철학 지식을 전달하는 길잡이 역할을 하며 학습자가 자신의 지식을 형성하는 과정에서 기존 지식의 재사용성을 극대화시킨다. 또한 학습자가 작성한 글을 개념화하여 자신의 온톨로지를 구축하고 이를 철학 온톨로지와 연계시킴으로써 발표 및 토론시에 비주열한 개념 전개를 보일 수 있고 원활한 토론과 객관적인 평가를 유도할 수 있다.

철학 온톨로지의 향후 과제는 철학 고전 텍스트의 지속적인 개념화와 함께 고전, 번역서 등의 전자 텍스트를 확보하여 이들을 철학 온톨로지와 연결함으로써 철학 지식 포털을 구현하는 것이다.

참 고 문 헌

- [1] 김도훈(2001). 디지털 시대의 인문학, 무엇을 할 것인가. 사회평론.
- [2] 홍윤기(2002). 철학콘텐츠 실태조사 및 수준 향상 방안. 인문정책연구총서 2002-47. 인문사회연구회/한국교육개발원.
- [3] Biezunski, M. Bryan, M. & Newcomb, S., ISO/IEC 13250 Topic Maps.
- [4] Dorbolo, Jon(2000). "The Philosopher's Web", In Academic Research on the Internet: Options for Scholars and Libraries. The Haworth Information Press.
- [5] Graham Moore. "Topic Map technology - the state of the art", XML 2000 Conference & Exposition, Washington, USA, December 2000.
- [6] Gruninger, M and Fox, M.S. "Methodology for the Design and Evaluation of Ontologies", IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, August 1995.
- [7] Lopez, M.F., Gomez-Perez, A., and Sierra, J.P. "Building a Chemical Ontology Using Methontology and the Ontology Design Environment", IEEE Intelligent Systems, 14(1), January 1999.
- [8] Mizoguchi, R. "Tutorial on ontological engineering-Part2: Ontology development, tools and languages", New Generation Computing, OhmSha&Springer, 22(1), 2004.
- [9] Noy, N.F. and McGuinness, D.L. "Ontology Development 101: A Guide to Creating Your First Ontology", SMI technical report SMI-2001-0880, 2001.
- [10] Online Papers in Philosophy, <http://opp.weatherson.net/>
- [11] Park, J. (2003). XML Topic Maps: Creating and Using Topic Maps for the Web. Pearson Education, Inc.
- [12] Perseus Classics Collection, http://www.perseus.tufts.edu/cache/perscoll_Greco-Roman.html
- [13] Philosophy in Cyberspace, <http://www-personal.monash.edu.au/~dey/phil/>
- [14] S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure. "Knowledge processes and ontologies", IEEE Intelligent Systems, Special Issue on Knowledge Management, 16(1), January/February 2001.
- [15] Stanford Encyclopedia of Philosophy, <http://plato.stanford.edu/>
- [16] Steve Pepper and Graham Moore. "XML Topic Maps(XTM) 1.0", TopicMaps.Org.
- [17] Uschold, M. and King, M. "Towards A Methodology for Building Ontologies", IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, August 1995.

저 자 소 개

정 현 숙(Hyun Sook Chung) 정회원

- 1993년 2월 : 대구가톨릭대학교 물리학과 (이학사)
- 1995년 2월 : 대구가톨릭대학교 전산학과 (이석사)



- 2003년 8월 : 연세대학교 컴퓨터과학과 (공학박사)
- 2004년 9월 ~ 현재 : 가톨릭대학교 컴퓨터정보공학부 교육전담교수
- <관심분야> : 캐릭터모델링, 모션그래픽스, 이터닝, 시맨틱웹

최 병 일(Byung-Il Choi) 정회원

- 1986년 2월 : 연세대학교 철학과 (문학사)
- 1988년 8월 : 연세대학교 철학과 (문학석사)



- 1997년 5월 : UC. Berkeley 철학과 논리학전공 (논리학박사)
- 2001년 9월 : 서울대학교 컴퓨터공학부 박사후 연구원
- 2002년 3월 ~ 현재 : 서울대학교 철학 사상연구소 전임연구원

<관심분야> : 논리체계, 지식표현, 온톨로지, 이터닝, 토픽맵