

클래스 시각화를 통한 OWL 온톨로지 편집기 개발¹⁾

김민수^{0*} 송세현* 서재우** 김민구**

아주대학교 정보통신전문대학원*, 아주대학교 정보통신대학 정보및컴퓨터공학부**
{bijey⁰, legoman, suhjwoo, minkoo}@ceai.ajou.ac.kr

Class-Visualization based OWL Ontology Editor

Minsoo KIM^{0*} Seheon SONG* Jaewoo SUH** Minkoo KIM**

Graduate School of Information and Communication Ajou University*
College of Information and Computer Engineering Ajou University**

요 약

시맨틱웹의 실용화에 대한 연구가 활기를 띠면서 국내에서도 시맨틱웹에 관한 연구와 그 기반이 되는 온톨로지에 대한 관심이 높아지고 있다. RDF, DAML+OIL 등 많은 온톨로지 언어가 개발되었고 W3C는 2004년 2월 OWL(Web Ontology Language)을 온톨로지를 위한 표준 언어로 발표했다. 이와 함께 프로티지 등 온톨로지 개발을 위한 편집 도구들에 관한 연구도 진행되고 있다. 본 논문에서는 기존 편집기의 문제를 지적하고 새로운 편집 방법을 통한 온톨로지 편집기를 개발하였다. 이 편집기는 가장 널리 사용되고 있는 온톨로지 개발도구인 프로티지와 그 위에서 작동하는 OWL plug-in을 기반으로 한다.

1. 서 론

차세대 웹이라 불리는 시맨틱웹은 인간만이 이해 할 수 있는 웹의 구조를 벗어나 기계에 의해 이해 가능한 웹의 구현을 목표로 하고 있다. 이미 시맨틱웹의 실용화에 대한 연구가 진행되고 있으며, 이제 더 이상 차세대 웹이 아닌 현재 작동하고 있는 웹이라는 주장도 제기되고 있다. 한편 시맨틱웹의 연구와 함께 그 기반이 되는 온톨로지에 대한 연구도 활기를 띠고 있다. W3C를 중심으로 OWL(Web Ontology Language) 등 다양한 온톨로지 언어들이 발표되었으며 온톨로지를 쉽게 편집하고 개발할 수 있는 개발 도구들에 관한 연구도 상당한 진전을 이루었다.

그러나 온톨로지 개발을 위한 언어들은 기계에 의한 의미 전달을 목표로 하고 있어 오히려 인간의 입장에서 이해하기 어려운 구조를 가지고 있다. 또한 온톨로지가 표현하고 있는 현상들과 그들 사이의 관계는 어려운 논리로 얽혀 있어 온톨로지 개발과 편집에 어려움이 있다. 이의 해결을 위해 온톨로지 편집을 쉽게 하기위한 기법에 관한 연구가 진행되고 있으며, 특히 온톨로지 시각화에 관한 연구는 온톨로지를 그래프 혹은 그림형태로 나타내준다는 점에서 흥미롭다. 그러나 이러한 연구에도 불구하고 여전히 온톨로지에 대한 직관적인 이해는 어려워 보이며, 규모가 큰 온톨로지의 경우에는 문제가 더욱 심각하다. 본 논문에서는 이러한 문제점들을 지적하고 온톨로지의 직관적인 이해와 편집을 위한 새로운 방법의 제시와 그에 기반한 편집기를 개발하였다. 이 편집기는 가장 널리 사용되고 있는 온톨로지 편집 도구인 프로티지와 그 위에서 작동하는 OWL plug-in을 기반으로 하고 있다.

2. 관련 연구

2.1 Ontology

온톨로지는 특정분야의 현상들에 대한 명확한 관계구명을 통한 개념화이며, 이는 기계에 의해 이해가능한 형태로 표현된다[1]. 온톨로지의 구성은 개발 언어, 개발 집단마다 용어의 차이는 있지만 기본적으로 클래스(class or concept), 속성(property, relation or slot)과 인스턴스(instance)로 이루어진다²⁾. 온톨로지는 시맨틱웹 위에서 서로 다른 어플리케이션 간에 공유되어 사용되는 지식기반의 역할을 하며, 시맨틱웹 혹은 시맨틱웹 서비스의 주요 이슈인 상호 운용성의 토대가 된다.[2]

2.2 OWL(Web Ontology Language)

W3C는 온톨로지를 기술하기 위한 언어로 RDF, DAML+OIL 등을 개발해왔고 2004년 2월 온톨로지 개발을 위한 언어의 표준으로 OWL을 발표했다[3]. XML(eXtensible Markup Language) 기반의 메타데이터 기술언어인 RDF(Resource Description Language)가 OWL의 문법(Syntax)을 제공하며 DAML+OIL 등은 OWL 시맨틱의 토대가 되었다. OWL은 기존 RDF(S), DAML+OIL 등에 비해 풍부한 표현력을 제공하기 위한 많은 특성들을 가지고 있어 온톨로지 개발에 있어 강력한 기능을 발휘한다. 그러나 기계가 이해할 수 있게 하는 강력한 표현력은 언어 구조의 복잡함과 논리적 어려움을 가져왔다. OWL 이하의 표현력을 가진 언어, 예를 들어 RDF(S)에 있어서도 이는 마찬가지이다. 따라서 현실 세계를 온톨로지 구성할 클래스, 속성 등을 추출해내는 것도 어려운 작업이지만 구성된 온톨로지를 OWL

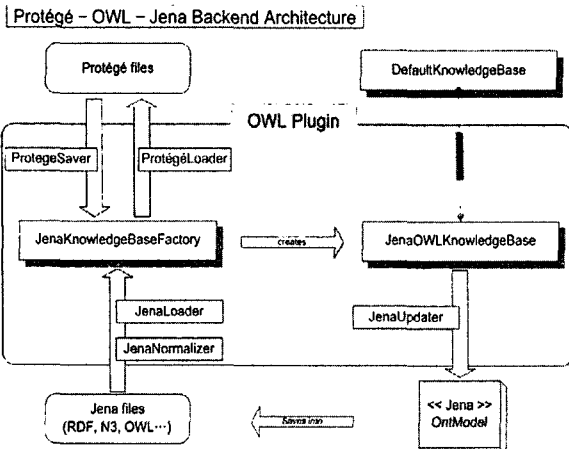
1) 본 논문은 과학기술부의 국가지정연구실 사업(과제명:차세대 인터넷을 위한 지능형 온톨로지 자동생성 시스템 개발, 과제번호:M10302000087-03J0000-04400) 지원으로 수행되었음

2) 본 논문에서는 OWL을 대상으로 하기 때문에 OWL의 키워드인 클래스(Class), 속성(property), 개체(instance)를 사용하기로 한다. 단 개체는 인스턴스로 표기한다.

등 온톨로지 언어를 통해서 표현하는 작업도 쉬운 일은 아닌 것으로 생각된다.

2.3 프로티지(Protege)

위에서 지적한 것처럼 온톨로지 표현 언어들의 복잡함을 해결하기 위해 OntoEdit, 프로티지 등 쉬운 온톨로지 개발과 편집을 위한 도구들이 개발되었다. 프로티지는 SMI(Stanford Medical Informatics)에서 개발한 Java 시스템 기반 온톨로지 및 KB 개발/관리 도구이다.[4] 1980년대 SMI 자체 지식관리 시스템인 ONCOCIN을 모태로 하고 있으며, 1990년대 프로티지 개발을 시작으로 현재 2.1.2 버전이 제공된다. 프로티지는 사용자에게 익숙한 인터페이스를 제공하고 소스가 공개되어 있어 온톨로지 편집을 위한 도구로 가장 널리 사용되고 있다. 또한 플러그인을 통한 프로티지 시스템의 확장은 프로티지가 갖는 가장 큰 장점이다. 현재 다양한 온톨로지 언어로의 import/export 기능, Biomedical Informatics, 온톨로지 시각화 기능, 데이터베이스와의 연동 기능 등을 구현한 플러그인이 50여 종류가 공개되어 있다.



<그림 1> Protege-OWL plug-in Architecture

본 논문에서 구현한 OWL 편집기의 기반이 되는 프로티지와 OWL plug-in의 구조는 <그림 1>과 같다. OWL plug-in은 Jena 분석기를 통해 OWL 온톨로지를 분석하여 프로티지 고유의 지식베이스와 동기화를 하며, OWL만이 가진 시맨틱을 편집하기 위한 추가기능, 추론 등의 기능을 제공하고 있다.

2.4 Ontology Visualization

프로티지 등 온톨로지 편집기들은 쉬운 방법으로 온톨로지를 개발할 수 있다는 데 의미가 있다. 그러나 온톨로지 구조의 파악, 직관적인 이해 등이 그리 용이하지 않다는 문제가 지적되었다. 이를 해결하기 위한 방법 중 온톨로지 시각화에 대한 연구가 진행되었고 OntoViz, ezOWL 등 프로티지 플러그인으로도 다수 개발되었다. 온톨로지 시각화에 관한 연구는 직관적인 온톨로지 이해와 쉬운 편집에 많은 도움이 되었다. 그러나 문제는 여전히 남아있다. 온톨로지를 시각적으로 보여주는 것은

사용자가 자신의 컴퓨터 화면을 통해서 볼 수 있는 범위에 한정되는 데, 일반적으로 온톨로지를 한 화면에 보여 주는 것은 무리가 있다. 클래스와 속성의 수가 늘어나게 되면 두 클래스간의 관계조차도 이해하기 어려워지고, 대규모 온톨로지의 경우에는 시각화의 의미가 무색해진다.3)

3. OWL Editor의 구현 및 결과

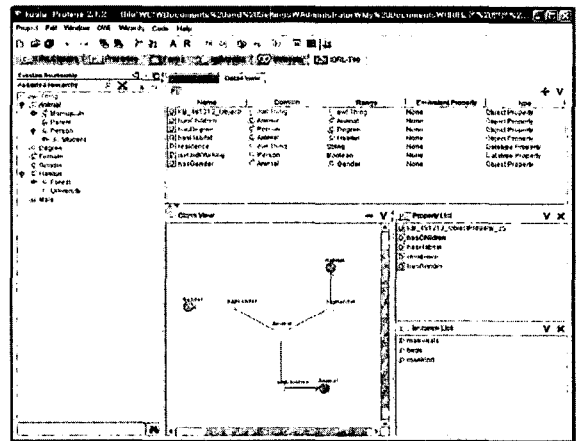
3.1 개발 동기 및 배경

OWL Editor는 OWL 문법의 어려움을 넘어 쉬운 OWL 온톨로지 편집을 목표로 하고 있다. 온톨로지 표현 언어로 OWL은 W3C의 권고안일 뿐 아니라 다른 언어보다 표현력이 풍부하기 때문에 OWL을 기반 언어로 선택하였다. 또한 세계의 많은 사용자들에 의해 사용되고 검증받고 있으며 지속적인 개발이 이루어지고 있는 프로티지를 구현의 기초로 삼았다.

본 논문에서는 온톨로지의 각 구성요소, 즉 클래스, 속성, 인스턴스를 사전작업을 통해서 추출한 것으로 가정하였다. 수작업 혹은 자동화에 의한 방법4)으로 추출된 구성요소들은 OWL Editor를 통해 온톨로지로의 형태를 갖추고 개발/관리 된다.

3.2 interface

OWL Editor는 <그림 2>에서 표현된 Class Tree, Property Pool, Class View의 세 부분으로 구성된다.



<그림 2> OWL Editor Interface

3.2.1 Class Tree(Hierarchy)

클래스 트리는 OWL 플러그인에서 제공하는 클래스 트리를 그대로 사용하였다[6]. 온톨로지의 구조는 클래스

3) 한 논문에서는 시맨틱 웹에서 동작하는 온톨로지를 소규모로 가정하였다[5]. 사실 소규모 온톨로지의 경우 시각화는 온톨로지 이해에 큰 도움이 된다.

4) 수작업에 의한 온톨로지 구축은 사실 지난한 작업일 뿐 아니라 규모가 큰 온톨로지 구축에 있어서는 바람직하지 못하다. 본인이 속한 연구실에서는 반자동 온톨로지 구축을 연구 중이고 본 논문에서 다루고 있는 OWL 편집기는 추출된 온톨로지 구성요소를 편집하고 관리하기 위함이다.

다중 상속 등으로 인하여 트리구조의 표현이 적절치 못한 면이 있는 바 이는 그래프 형태로의 온톨로지 시각화의 이유이기도 하다. 그러나 클래스의 상속관계는 온톨로지를 파악하는 데 중요한 열쇠이고 그래프 형태에서 상속관계(subclass) 자칫 간과될 수 있으며 그래프를 복잡하게 만드는 이유가 되어 그래프에서 생략하고 클래스 트리를 삽입하였다. 클래스 트리에서는 클래스 선택, 클래스 상세보기, 클래스 추가/삭제 등의 기능이 수행된다.

3.2.2 Property Pool

Property Pool은 온톨로지 내에 존재하는 모든 속성 목록을 가지며 단순보기(Simple View)와 자세히 보기(Detail View) 두 형태를 지원한다. Simple View는 단순히 속성의 이름만을 보여주고, Detail View는 속성의 이름, 도메인(domain), 범위(Range), 형식(type) 동등 속성(equivalent property) 등 속성의 정보를 테이블 형태로 보여준다. Detail View에서 속성의 정보를 표시한 이유는 각 속성에 대한 정보를 사용자가 기억하고 있을 수 없고 새로운 창에서 확인하는 것은 번거로운 일이기 때문에 클래스에 적당한 속성을 바로 적용하기 위함이다. Property Pool과 아래에서 설명할 Class View는 마우스 드래그 앤 드랍이 지원되어 마치 MS 윈도우즈 탐색기처럼 속성을 원하는 클래스에 쉽게 적용할 수 있다.

3.2.3. Class View

Class View는 클래스 트리에서 선택한 클래스를 그래프 형태로 보여준다. 온톨로지 시각화 연구에서 지적한 대로 온톨로지가 커질수록 시각화는 무의미해진다. 따라서 온톨로지 전체에 대한 시각화보다는 선택된 클래스에 대한 시각화에 중점을 두어 선택된 클래스를 그래프의 중앙에 위치시키고 속성으로 연관된 클래스들을 바깥에 위치시키는 구조를 선택하였다. 이 구조는 한 클래스에 대한 자세한 정보를 직관적으로 확인할 수 있다. 그래프에서 한 속성으로 연결된 두 클래스를 고려하면 RDF Triple, 즉 S-P-O(Subject-Predicate-Object)의 구조이다. 또한 OWL이 가지고 있는 Property Constraint 등을 그래프에 표시하였다. Property Pool에서 설명한대로 마우스 드랍을 지원하여 드래그된 속성을 그래프에 드랍하면 클래스에 속성이 추가되고 온톨로지에 즉각 반영된다. 클래스 혹은 속성에 대한 상세정보 보기가 가능하고 세부항목 등을 수정할 수 있다. Class View는 그래프 외에도 선택 클래스가 가지는 속성, 인스턴스 리스트를 따로 보여주는 데 이는 그래프만으로 부족한 정보를 표시하기 위함이다.

4. 결과 및 향후 과제

본 논문에서 다룬 OWL 온톨로지 편집기는 기존의 시각화 방법에서 벗어난 새로운 시각화 방법을 제시하고 구현하였다. 이는 선택한 클래스 중심의 그래프를 보여주며, 그 클래스를 편집하는 관점에서 필요한 정보들을 보여주고 쉬운 방법의 편집이 가능하도록 구현되었다. Property Pool에서 제공되는 속성들은 굳이 속성 각각을 살펴볼지 않더라도 속성의 일반적인 특성을 알 수 있게

해준다. 또한 드래그 앤 드랍 형식의 클래스 편집은 윈도우 환경에 익숙한 사용자에게 직관적으로 다가갈 수 있어 긍정적으로 평가된다.

그러나 OWL 편집기에서 그래프 부분은 온톨로지를 시각적으로 보여주기엔 불충분하다. 이는 대규모 온톨로지 시각화에 대한 반평향으로 단지 하나의 클래스만을 보여주는 단점으로 작용할 수 있다. 따라서 선택된 클래스로부터 n-거리에 위치한 부분을 보여주는 등의 방법적인 연구가 필요하다. 또한 현재의 편집기는 오로지 프로티지에서 제공하는 그래픽 인터페이스만으로 온톨로지를 보여주고 편집한다. 그러나 클래스, 속성 등에 대한 정보 파악은 오히려 OIL 등 프레임 베이스 언어가 직관적 이해에 도움이 될 것으로 보인다. 따라서 인간에게 익숙한 언어로의 정보 표현과 이를 위한 편집 기술의 구현도 온톨로지의 쉬운 편집을 위한 향후 과제로 남는다.

5. 결론

온톨로지 개발에 있어서 기계에 의해 이해 가능한 형태인 온톨로지 언어는 온톨로지 개발에 어려움으로 작용하기 때문에 쉬운 온톨로지 개발도구의 구현이 필수적이다. 또한 개발도구는 쉬운 편집 뿐 아니라 사용자에게 온톨로지에 대한 직관적인 이해를 제공해야 한다.

본 연구에서는 온톨로지 언어의 표준으로 자리 잡은 OWL과 가장 널리 사용되고 있는 온톨로지 편집기 프로티지 및 OWL 플러그인을 기반으로 새로운 방법의 OWL 온톨로지 편집기를 구현하였다. OWL 온톨로지 편집기는 온톨로지의 속성들에 대한 정보를 한눈에 보여주고 이를 클래스에 쉽게 적용할 수 있게 한다. 또한 전체 온톨로지에 대한 대략적인 시각화를 제공하는 것이 아니라 그래프 형태를 통해서 특정 클래스를 중심으로 상세한 시각화를 제공한다. 특히 이러한 방식은 온톨로지의 규모가 커질수록 효과적일 것으로 판단된다. 나아가 그래프의 성능 향상 및 프레임 기반 언어와 편집에 대한 연구가 더해진다면 OWL 온톨로지 편집기는 OWL 온톨로지의 효율적인 개발 및 편집에 기여할 것이다.

6. 참고문헌

- [1] Thomas R. Gruber, "Toward Principles for the Design of Ontologies Used for Knowledge Sharing", 1993
- [2] 최중민, "Semantic Web : Overviews and Trends", 한국어정보처리학회 2004 동계 튜토리얼, p35-77
- [3] <http://www.w3.org/TR/>
- [4] <http://protege.stanford.edu>
- [5] Christiaan Fluit, Marta Sabou and Frank van Harmelen, "Supporting User Tasks through Visualization of Light-weight Ontologies", Handbook on Ontologies, p425-432
- [6] <http://protege.stanford.edu/plugins/owl/index.html>
- [7] Oscar Corcho, "A Roadmap to Ontology Specification Languages", 2000