

개체 시각화 중심의 OWL 온톨로지 편집기 개발¹⁾

김민수^o* 김준영* 송세현* 김민구**

아주대학교 정보통신전문대학원*, 아주대학교 정보통신대학 정보및컴퓨터공학부**
{bijey^o, tristan, legoman, minkoo}@ceai.ajou.ac.kr

Element-Visualization based OWL Ontology Editor

Minsoo KIM^o* Junyoung KIM, Seheon SONG* Minkoo KIM**

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

요 약

시맨틱 웹과 시맨틱 웹 서비스의 논리적 기반이 되는 온톨로지에 대한 연구가 활발히 진행되고 있다. W3C는 웹 온톨로지 기술을 위한 표준 언어로 OWL을 발표했고, OWL로 기술된 온톨로지는 웹 서비스, 유비쿼터스 컴퓨팅 등에서 활용되고 있다. 그러나 OWL의 복잡한 문법 구조는 OWL 온톨로지 개발을 어렵게 하고 있어, 직관적으로 OWL 온톨로지를 개발할 수 있는 방법이 요구되었고, 이를 위해 Protege 등 많은 온톨로지 편집기들이 개발되었다. 본 논문에서는 쉬운 OWL 온톨로지 편집을 위해 개체 중심의 온톨로지 시각화를 통한 온톨로지 편집기 개발에 관한 내용을 담고 있다. 이 편집기는 JDK 1.4 환경에서 개발되었으며 OWL 온톨로지 분석을 위해 Jena2 API를 사용하였다.

1. 서 론

기계에 의해 이해 가능한 웹을 목표로 하고 있는 시맨틱 웹과 웹 서비스에 대한 연구가 활발히 진행되고 있는데 그 논리적 기반이 되는 온톨로지에 관한 연구와 활용이 주목받고 있다. W3C에서는 웹 온톨로지를 기술하기 위한 표준으로 OWL(Web Ontology Language)을 발표하였고, OWL로 개발된 온톨로지들이 다양한 분야에서 활용되고 있다. 그러나 OWL의 복잡한 문법 구조는 OWL 온톨로지 개발에 어려움을 주고 있다.

이의 해결을 위해 많은 연구들이 진행되었고, Jena, OWL API 등 OWL 온톨로지에 쉽게 접근할 수 있는 도구들이 개발되었다. 현재 가장 널리 사용되고 있는 온톨로지 편집기인 Protege의 OWL 플러그인 역시 Jena를 이용하여 OWL 온톨로지 개발 및 편집을 가능하게 하고 있다. 또한 ezOWL 등 복잡한 구조의 온톨로지를 직관적으로 이해할 수 있도록 하는 온톨로지 시각화에 관한 연구도 성과를 나타내었다.

그러나 이러한 연구에도 불구하고 OWL 온톨로지의 직관적인 파악은 미흡해 보이며, 특히 온톨로지 전체에 대한 시각화에 초점을 맞춘 온톨로지 시각화는 특정 개체의 이해 및 규모가 큰 온톨로지 시각화에는 문제가 있어 보인다. 본 논문에서는 이러한 문제점들을 지적하고 개체중심 시각화를 통한 OWL 온톨로지 편집기를 개발하였다. 이 편집기는 JDK 1.4 환경에서 개발되었고 OWL 온톨로지를 다루기 위해 Jena2 API를 이용하였다.

2. 관련연구

2.1 OWL(Web Ontology Language)

의미를 담은 웹의 표현과 기계에 의한 의미의 획득을 가능하게 하기 위하여 많은 언어가 개발되었고, W3C는 2004년 2월 OWL을 웹 온톨로지 개발을 위한 표준으로 발표하였다. 온톨로지 언어에 관한 연구는 시맨틱 웹에 문법적 구조를 제공하는 XML(eXtensible Markup Language)을 시작으로, RDF(Resource Description Language) / RDFS로 이어졌다. RDF(S)는 웹에 의미를 부여할 수 있게 하였지만, 표현 능력이 부족하다고 평가되었고 이를 보완하여 OIL, DAML, DAML+OIL 등이 개발되었다. OWL[4]은 DAML+OIL의 시맨틱을 보완하여 개발되었고, RDF(S), DAML+OIL 등에 비해 풍부한 표현력을 제공하기 위한 많은 특성들을 가지고 있어 온톨로지 개발에 있어 강력한 기능을 발휘한다.

그러나 기계가 이해 할 수 있는 강력한 표현 능력의 구현은 언어 구조의 복잡함과 논리적 어려움을 가져왔다. 즉 기계에 의해 이해 가능한 언어의 기술이 오히려 사람에게는 익숙하지 못한 언어가 되었다. 따라서 다른 개발 도구의 도움 없이 OWL 온톨로지를 개발하는 것은 매우 어려워 보인다.

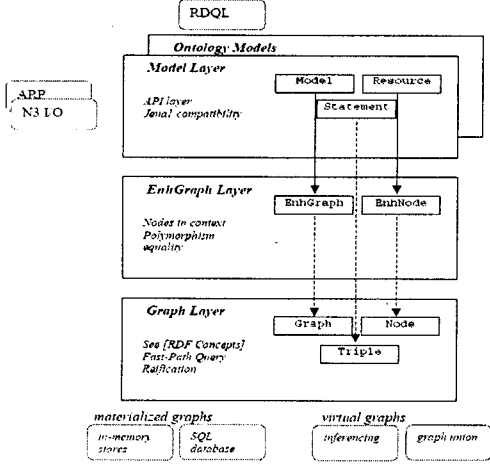
2.2 Jena

Jena는 휴렛 팩커드사가 개발한 시맨틱 웹 서비스 혹은 어플리케이션을 위한 자바로 개발된 프레임워크이다. 2000년 Jena1, 2003년 Jena2를 발표하였으며 시맨틱 웹 및 온톨로지 어플리케이션에서 널리 사용되고 있다. Jena는 RDF(S), OWL을 위한 프로그램 환경, 룰 기반 추론 엔진 등을 지원한다. Jena2의 주요 기능은 RDF API, OWL API, In-memory and persistent storage(Database), RDQL - a query language for RDF 등이다.

본 논문에서 구현한 OWL 편집기도 Jena의 OWL API를 이용한다. Jena는 기본적으로 RDF Graph(S-P-O triple)

- 1) 본 논문은 산업자원부의 국가지정연구실 사업(과제명:차세대 인터넷을 위한 지능형 온톨로지 자동생성 시스템 개발, 과제번호:M10302000087-03J0000-04400) 지원으로 수행되었음.
- 2) 본 논문에서 말하는 개체는 온톨로지에 포함된 클래스, 프로퍼티 등 온톨로지의 모든 구성요소이다.

를 기반으로 온톨로지 모델을 표현한다. 그런데 OWL은 RDF의 문법에 기반을 두지만 RDF를 넘어서는 특징을 가지기 때문에, Jena는 OWL 온톨로지를 단순한 RDF serialization이 아니라 RDF-centric view로서 다룬다.[3] Jena는 OWL Full³⁾에 특성까지 Model로서 유지하며, 사용자는 어플리케이션 내에서 이 Model을 이용하여 OWL 온톨로지 개발 및 편집이 가능하다.



<그림 1> Jena2 Architecture

2.3 Ontology Editing and Visualization

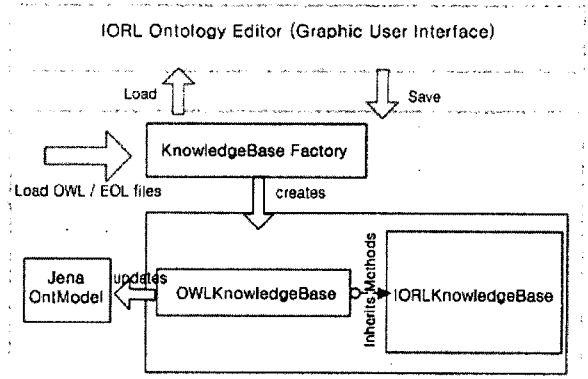
온톨로지 시각화는 복잡한 구조의 온톨로지 혹은 OWL 등 태깅 언어로 작성된 온톨로지를 쉽게 이해하고 편집할 수 있게 한다. Protege⁴⁾, OntoEdit⁵⁾ 등 대표적인 온톨로지 편집기들은 그래프, 트리 등을 이용한 온톨로지 시각화를 제공하며 ETRI에서 개발한 ezOWL⁶⁾은 OWL 온톨로지 시각화 도구로 좋은 평가를 받고 있다.

그러나 대부분의 온톨로지 시각화 도구는 온톨로지에 포함된 전체 개체를 대상으로 한다. 따라서 개체의 수가 증가함에 따라 그래프는 복잡해지고 관계와 의미의 파악이 어려워진다. 이를 테면 고정된 화면에 많은 개체가 표현이 되려면 개체가 작은 크기로 표현이 되거나 멀리 떨어진 개체와의 관계 표현이 어려워진다.

3. OWL Ontology Editor

3.1 Overview and Architecture

본 논문에서 구현한 온톨로지 편집기는 전체 온톨로지가 아닌 현재 고려되고 있는 개체를 중심으로 한 시각화를 통해 OWL 온톨로지의 쉬운 편집을 목표로 개발되었다. 온톨로지 편집기는 JDK 1.4 환경에서 개발되었으며, OWL 온톨로지 분석을 위해 Jena OWL API를, 그래프를 통한 시각화에는 JGraph 라이브러리를 사용하였다. <그림 2>는 온톨로지 편집기의 Architecture를 보여주고 있다.7)



<그림 2> 온톨로지 편집기 Architecture

3.2 Interface - Basic Features

3.2.1. Class Tree

클래스 트리는 온톨로지에 포함된 클래스의 Hierarchy를 보여준다. 온톨로지는 완벽하게 트리 구조에 부합되는 것은 아니지만, 기본적으로 온톨로지의 구성은 Hierarchy를 중심으로 구성된다는 면에서 트리구조는 유용하다. 클래스 트리에서는 클래스의 생성, 제거, 이름의 수정 등 간단한 편집이 가능하다.

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.

3.2.3 Graph Panel

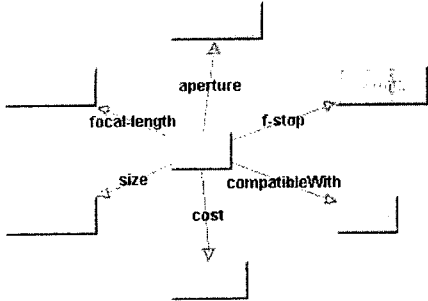
Graph Panel은 클래스 트리에서 선택된 클래스 혹은 Property Pool에서 선택된 프로퍼티를 중심으로 그래프를 그린다. 위에서 지적했듯이 전체 온톨로지를 대상으로 하는 시각화는 온톨로지의 크기가 커질수록 무의미해진다. 따라서 온톨로지 전체에 대한 시각화보다는 선택된 클래스 혹은 프로퍼티에 대한 시각화에 중점을 두어 선택된 클래스를 그래프의 중앙에 위치시키고 속성으로 연관된 클래스들을 바깥에 위치시키는 구조를 선택하였다. 이 구조는 임의의 개체에 대한 직관적인 이해에 효과적이다. 아래 그림은 OWL로 기술된 camera 온톨로지에서도 Lens 클래스를 그래프로 표현한 것이다. 노란색 노드는 사용자 정의 클래스를, 파란색 노드는 XML schema 데이터 타입을, 예지는 property를 나타내면 예지 이름은 property 이름이다.

한편 OWL의 Equivalent Class, Property Constraint

3) OWL은 Light, DL, Full의 세 버전으로 발표되었다.[4]
 4) http://protege.stanford.edu/
 5) http://www.ontoknowledge.org/tools/ontoedit.shtml
 6) http://iweb.etri.re.kr/ezowl/

7) OWL 뿐 아니라 다른 언어 지원을 위해 KnowledgeBase의 설계를 그림과 같이 하였다.

등은 복잡한 구조로 이해되기 어렵기 때문에 자세한 그 래프로 표현하여 쉬운 이해를 가능하게 했다.



<그림 3> Class 시각화 예제

또한 전체 온톨로지의 구조만을 보여주는 네비게이션 창이 존재한다. 즉 본 논문에서는 전체 온톨로지에 대한 구조를 보여주되, 개체에 초점을 맞춘 시각화 전략을 선택했다.

3.2.4 Source Viewer

Source Viewer에는 현재 고려되고 있는 개체에 대한 OWL Source를 보여준다. 온톨로지 개발/편집에 있어 OWL Source를 참고하는 것은 종종 유용하다.

3.2.5 Detail Information Window

각 개체에 대한 상세한 정보를 새로운 윈도우를 통해 보여준다. 클래스 트리나 Property Pool에서 편집할 수 없는 상세한 편집기능을 제공하며, 이 윈도우에도 역시 그래프가 포함된다. 여기에 표현되는 그래프는 편집하는 대상만을 표현하는 간단한 형태이다.

4. 결과 및 향후 과제

본 논문에서는 개체 중심의 온톨로지 시각화를 이용한 편집기를 개발하였다. 이는 고려하고 있는 개체를 중심으로 그래프를 보여주며 그 개체에 대한 직관적인 이해를 가능하게 한다. 또한 사용자에게 익숙한 인터페이스 및 쉬운 편집기능은 어려운 OWL 온톨로지 편집에 직관적으로 다가갈 수 있어 긍정적으로 평가된다.

한편 이에 더하여 온톨로지 편집기에 다음과 같은 확장된 기능들이 필요할 것으로 생각되며 현재 개발 단계에 있다.

4.1 Persistent Storage - Database

대용량 온톨로지를 저장하기 위한 데이터베이스 스키마를 개발하여 데이터베이스를 이용한 온톨로지 열기/저장 기능을 지원한다.

4.2 Epistemological Ontology Language (EOL)

OWL의 복잡한 문법 대신 프레임 기반 언어가 온톨로지의 이해에 도움이 될 것이다. 따라서 인간에게 익숙한 언어로의 표현이 온톨로지 개발에 도움이 될 것이다. EOL은 OWL로 변환 될 수 있다.

4.3 Automatic Ontology Extraction

문서 집합에서 클래스와 프로퍼티를 자동으로 추출하는 모듈이다. 온톨로지 자동 구축은 매우 어려운 과제로 보이

지만, 온톨로지 개발을 훨씬 쉽고 빠르게 할 수 있다.

| Modeling Features | Represent Ontology with Element based Visualization |
|------------------------|---|
| Base Language | OWL, EOL * |
| Import/export | OWL, EOL*, Database* |
| Graph View | Browsing Elements, Navigation for whole Ontology |
| Consistency Checks | Not Yet |
| Information Extraction | Semi-Automatic Extraction for Class and property (relation) from documents* |
| Ontology Merging | Not Yet |

* : 개발중

<표 1> 온톨로지 편집기 요약

5. 결론

시맨틱 웹의 논리적 기반이 되는 온톨로지는 그 구조에 있어서 혹은 표현하는 언어에 있어서 복잡한 구조를 가지고 있다. 따라서 온톨로지 개발에 있어서 이러한 어려움을 해소하는 온톨로지 편집기는 필수적인 요소이다. 또한 온톨로지 편집기는 사용자에게 온톨로지의 쉬운 편집과 직관적인 이해를 다양한 방식으로 제공해야 한다.

본 연구에서는 개체 중심의 시각화를 중심으로 OWL 온톨로지에 대한 쉬운 접근을 가능하게 하였다. 기존의 전체 온톨로지에 초점을 맞춘 시각화에서 벗어나 각 개체를 시각화의 중심에 위치시킴으로서 현재 고려하고 있는 개체에 대한 직관적인 이해를 가능하게 한다. 또한 간단한 그래프를 통한 개체의 편집, Property Pool에서의 프로퍼티 정보 표현 등도 쉬운 OWL 온톨로지 개발을 가능하게 한다.

나아가 큰 규모의 온톨로지를 저장할 수 있는 데이터베이스의 이용, 인간에게 익숙한 온톨로지 언어의 개발 및 활용, 그리고 온톨로지 자동 구축 모듈에 대한 연구가 추가된다면 온톨로지 편집기는 더욱 강력한 기능을 발휘한 것이며, OWL 온톨로지 개발 및 편집을 자유롭게 할 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", 한국여정정보처리연구회 2004 동계 튜토리얼, p35-77
- [3] Jeremy J. Carroll, Ian Dickinson, Chris Dollin, Jena:Implementing the semantic web Recommendations, 2003, Hewlett-Packard Company
- [4] http://www.w3.org/2004/OWL/
- [5] Christiaan Fluit, Marta Sabou and Frank van Harmelen, "Supporting User Tasks through Visualization of Light-weight Ontologies", Handbook on Ontologies, p425-432
- [6] Oscar Corcho, "A Roadmap to Ontology Specification Languages", 2000