

상품 온톨로지 모델링 도구의 설계

탁문희^o 김경화 심준호

숙명여자대학교 컴퓨터학과

{mhtark^o, kamza81, jshim}@sookmyung.ac.kr

Design of a Modeling Editor for Product Ontology

Moonhee Tark^o, Kyunghwa Kim, Junho Shim

Dept. of Computer Science, Sookmyung Women's University

요약

시맨틱웹의 목적은 메타데이터의 개념을 적용하여 웹문서에 시맨틱 정보를 덧붙여, 웹에 있는 정보를 컴퓨터가 이해할 수 있도록 하는 것이다. 이를 가능하게 해주는 핵심 기술로 웹 온톨로지 언어 (OWL : Ontology Web Language)를 들 수 있다. 그러나 온톨로지 언어가 표현하는 객체들과 그들 사이의 관계는 복잡한 논리들로 얽혀 있어 언어에 익숙하지 않은 사람에게는 온톨로지 개발이 쉽지 않다. 본 논문은 개념 모델링을 위한 방법으로 가장 잘 알려진 EER 모델링만으로 복잡한 논리적 구조로 인해 작성하기 어려운 OWL 코드를 자동 생성하는 편집 도구를 설계한다. 특히 전자 카탈로그 도메인에 특화하여 자주 요구되는 특수한 개념들에 대한 모델링 요소를 추가하여 표현할 수 있게 하였다. 이는 온톨로지 개발자들의 시간과 어려움을 크게 감소시킬 것이며, 상품온톨로지 제작에 특히 유용할 것이다.

1. 서론

HTML 형태의 문서들로 이뤄진 현재의 웹은 인간에게 정보를 주는 중요한 역할을 하고 있지만, 컴퓨터 프로그램이 각 문서의 내용을 알아서 정확하게 파악 할 수 없다는 문제점을 갖고 있다. 이러한 문제의식에서 출발한 시맨틱웹 (Semantic Web)은 컴퓨터가 인터넷상의 웹문서의 내용을 이해하고 자동으로 처리하여 지식화 함으로써, 정보의 효율적 검색, 통합, 재사용을 가능하게 하는 지식의 근원으로서의 웹을 만드는 데 목표를 둔다. 시맨틱웹의 실용화에 대하여 이미 많은 연구가 진행 중이며, 이 기술의 기반이자 핵심인 온톨로지 (Ontology) 또한 크게 주목 받고 있다. 온톨로지는 특정분야의 현상들에 대한 명확한 관계규명을 통한 개념화이며, 이는 기계가 이해 할 수 있는 형태로 표현된다. W3C는 온톨로지를 기술하기 위한 다양한 온톨로지 언어들을 개발해왔고, OWL (Web Ontology Language)을 온톨로지 표준 언어로 발표했다. OWL은 다른 언어들에 비해 풍부한 표현력을 제공하지만, 인간이 추상적으로 바로 연상해 낼 수 있는 단순한 의미들까지 모두 정의해 주어야 기계해석단계 (machine level)에서의 추론이 가능하기에 오히려 인간의 입장에서 이해하기 어려운 구조를 가지고 있다. 또한 온톨로지 언어가 표현하는 객체들과 그들 사이의 관계가 복잡한 논리들로 얽혀 있어 온톨로지 언어에 익숙하지 않은 사람에게는 온톨로지 개발이 쉽지 않다. 최근 이러한 어려움을 해결하기 위한 연구가 활발히 진행되고 있으며 특히, 직관적인 이해를 돕기 위해 온톨로지를 그래프 혹은 그림의 형태로 나타내는 온톨로지 시각화를 위한 시도가 이어지고 있다. SMI (Stanford

Medical Informatics)에서 제작한 Protégé[4]와 i*COM에서 제작한 i*com Tool[2]은 온톨로지 시각화 편집도구의 예로써 현재 많은 온톨로지 연구가들에 의해 사용되고 있다.

이에, 우리의 시스템은 온톨로지 개발을 도울 수 있는 온톨로지 편집도구의 기능에 더하여, 차별화 된 다음과 같은 기능을 가지고 있는 편집도구를 설계하고자 하였다. 첫째, Description Logic 표현방식과 FacT 추론엔진 이 서로 Tightly-Coupled 한 i*COM과 달리 우리의 편집 도구는 OWL 코드를 생성하고, 이를 지원하는 추론엔진과 Loosely-Coupled 하다. 둘째, 온톨로지를 정의하는 과정에 필요한 EER (Extended Entity Relationship) 모델을 작성하면, 정의된 관계들을 OWL 코드로 변환해 주고, 생성된 OWL 코드의 직접 수정도 지원한다. 마지막으로, 선행연구[1]로 진행된 'e-Catalog의 온톨로지 모델링 기법'에서 제안되었던 전자카탈로그 (e-Catalog) 도메인에 특화된 모델링 요소 (Modeling Construct)를 추가하였다. 즉, 범용 도메인에서의 온톨로지 표현 뿐 아니라 공동구매(PurchaseSet), 대체품(Substitute), 보완품(Complement)과 같은 전자카탈로그 도메인에서 자주 요구되는 개념 및 관계들에 대한 새로운 모델링 요소를 제공한다. 본 논문에서는 이러한 차별화된 기능 실현을 위한 방법을 제안하고, 도구의 설계과정과 구현되어질 실행화면을 살펴보고 결론 및 향후 과제를 논한다.

2. 관련연구

2.1 상품 온톨로지 모델

전자카탈로그 도메인에서 모델링 된 상품 온톨로지는 컨셉 (Concept)과 관계 (Relationship)에 의해 정의 된다. 이는 4가지의 주요컨셉 (Key Concept : 상품

1) 본 연구는 정보통신부의 IT연구센터(ITRC)지원을 받아 수행되었음.

(Products), 분류체계(Classification Scheme), 상품속성 (Attributes), 단위(UOM)), 일반적인 의미 관계 (General Relationships), 전자카탈로그 도메인에서 추가적으로 요구되는 몇 가지 특별한 의미 (Specific Relationships) 관계를 포함한다.[1] 일반적 도메인에서 사용되는 의미 관계로는 ISA 관계, 부분체관계 (Meronymic Inclusion), 속성 및 유의어관계 등이 있다. 반면, 전자상거래 및 전자카탈로그 도메인에 특화하여, 고려될 수 있는 의미적 관계로는 공동구매 (PurchaseSet), 대체품 (Substitute), 보완품 (Complement) 등이 있다. 이러한 관계들은 상품 온톨로지 개발에 있어 유용하고, 필수적이다.

추가된 의미적 관계	의미
연필 - 대체품 - 볼펜	연필은 볼펜의 대체품이다.
충전지 - 보완품 - 충전기	충전지는 충전지의 보완품이다.
차량용거치대 - PDA - 공동구매 - 확장메모리 - 액정보호필름 - 네비게이션	PDA와 공동구매 관계가 있는 것은 액정보호필름, 확장메모리, 차량용거치대, 네비게이션이다.

<표 1> 상품 온톨로지에서의 특수 관계의 예 : 대체품, 보완품, 공동구매

2.2 Ontology Tool

2.2.1 Protégé

프로티지는 시맨틱웹 언어를 사용한 개념모델링과 온톨로지 편집 및 구축을 위한 시각적 도구이다. SMIE에 의해 Java 기반으로 개발되었으며, 온톨로지의 개발, 관리를 위한 지식 기반 시스템으로서 소스가 공개되어 있는 오픈 소스 개발 환경이다. 1990년대 프로티지 개발을 시작하여 현재 3.1 Beta 버전이 배포되고 있으며, 사용자에게 익숙한 그래픽 인터페이스를 제공하고, 또한 각종 플러그인을 통하여 시스템을 유연성 있게 확장시킬 수 있다. 현재 다양한 온톨로지 언어의 Import, Export 기능, Biomedical Informatics, 온톨로지 시각화기능, 데이터베이스와의 연동기능 등을 구현한 플러그인 50여 종류가 공개되어 있다. 프로티지의 이러한 유연하고 확장 가능한 지식모델과 오픈플러그인 구조는 시맨틱웹 언어를 위한 여러 개념레벨 편집기 개발의 기반이 되었다.

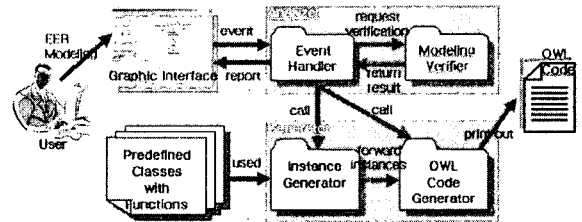
2.2.2 I*COM

ICOM은 사용자가 Intra-schema Constraint 를 사용하여 EER 모델링을 가능하게 하는 고급 CASE Tool 이다. 이 도구는 일관성 검사와 추론을 위해 매우 강력한 DL (Description Logic) 추론서버와 완전하게 통합되어 있다. ICOM은 EER 개념모델링을 위하여 처음으로 구현된 도구이다. 본 논문에서 구현된 도구처럼 EER 개념 모델을 OWL 코드로 변환하는 기능은 지원되지 않지만, 사용자에게 친숙한 그래픽 인터페이스를 통하여 몇 가지

의 연결된 개념 스키마를 생성, 편집, 관리, 저장 할 수 있다. 표준 Java 1.2에서 작성되어 리눅스와 윈도우 머신에서 사용되고 있는 ICOM은 CORBA Protocol 을 통하여 FaCT DL 서버와 통신하고, XML 포맷에서의 스키마와 Rational Rose™ 다이어그램을 Importing, Exporting 하기 위한 인터페이스를 제공한다.

3. 시스템 설계

우리가 구현한 온톨로지 편집 도구는 사용자가 그래픽 도구를 이용하여 설계한 EER 모델의 관계들을 분석 및 해석하여 OWL 코드를 생성한다.



<그림 1> 상품 온톨로지 모델링 도구의 구성도

- Graphic Interface : 사용자가 그래픽 도구를 사용하여 EER 모델링을 한다. 이때, 상품 온톨로지에 추가된 특별한 의미 관계를 적용 할 수 있다.
- Analyzer
 - Event Handler : 다이어그램이 자유롭게 그려질 때마다 발생하는 이벤트를 처리하고 설계된 모델의 개체와 연결 관계를 분석하기 위해 Modeling Verifier 을 호출한다.
 - Modeling Verifier : EER 모델의 타당성을 검증한다.
- Generator
 - Instance Generator : 모델로 정의된 객체들의 인스턴스를 생성한다.
 - OWL Code Generator : 코드생성 이벤트가 발생하면 생성된 인스턴스를 확인하여 OWL 코드를 생성한다.
- Predefined Classes with Functions : EER 모델링의 결과를 JAVA 인스턴스화 하고, OWL 코드로 변환시키기 위해 미리 정의해 놓은 클래스와 함수들로서, 각 객체의 Instance 생성 시 사용된다.

4. 시스템 구현

우리는 Windows 2000 Professional 운영체제에서 Java 2 Platform Standard Edition 5.0과 JBuilder 를 사용하여 온톨로지 편집 도구를 개발한다.

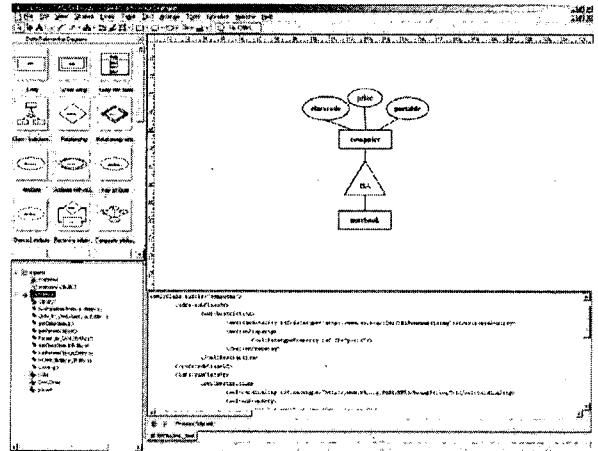
프로그램 코드 작성에 앞서 EER 다이어그램 개체들의 관계와 이것의 OWL 표현 간의 상호 의미관계 정립 과정 및 매핑 과정을 선행하였다. 이를 구현하기 위하여 EER 다이어그램 요소 (Entity, Attribute, ISA Relationship, etc) 별로 클래스를 정의하고 그려지는 개

체마다 인스턴스를 생성하여 상태와 관계를 유지하도록 하였다. 아래의 <표2> 는 EER 다이어그램 요소별로 정의된 대표적인 Java Class 들을 표시하고, 컴퓨터와 노트북의 ISA Relationship 을 예로 들어 EER 다이어그램과 OWL 의 표현을 간략히 비교하여 보여주고 있다.

EER	Java Class
	<pre>public class Entity extends OBJECT{.... public void appendAttribute(Attribute o) public int getAttributeSize().... } public class ISARelation extends OBJECT{.... public void setParentObject(Entity o) public void setChildObject(Entity o).... } public class Attribute extends OBJECT{.... public void setAttributeName(String n) public void setAttributeType(String t).... }</pre>
OWL Code	
<pre><owl:Class rdf:ID="computer"> <rdfs:subClassOf> <owl:Restriction <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#String">1</owl:cardinality> <owl:onProperty> <owl:DatatypeProperty rdf:ID="price"/> </owl:onProperty> </owl:Restriction> </rdfs:subClassOf> </owl:Class> <owl:Class rdf:ID="notebook"> <rdfs:subClassOf rdf:resource="#computer"/> </owl:Class> <owl:DatatypeProperty rdf:about="#price"> <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#String"/> <rdfs:domain rdf:resource="#computer"/> </owl:DatatypeProperty></pre>	

<표 2> EER, Java Classes, OWL Mapping 예제

아래의 <그림2> 는 구현된 온톨로지 편집 도구의 실행화면 구성을 보여준다. EER 모델링을 위한 그리기 도구 모음과 다이어그램을 편집하는 프레임, 그리고 모델을 JAVA 인스턴스로 생성하며, 분석되는 객체 메소드를 트리 형태로 보여주는 프레임과 EER 모델을 OWL 코드로 변환, 생성하여 보여주는 결과 코드 출력 프레임으로 구성되어 있다.



<그림 2> 상품 온톨로지 모델링 도구 실행화면

5. 결론 및 향후연구

본 연구에서는, 전자카탈로그 도메인에서 자주 요구되는 특수한 개념들에 대한 모델링 요소를 새로 추가하여 제공하며, 이것을 사용한 모델링 결과를 OWL 코드로 자동 변환 해주는 시각적 온톨로지 개발 편집 도구를 설계하였다. 이러한 설계에 기반 하여 개발된 도구는 온톨로지 개발에 드는 시간을 크게 감소시킬 것이며, 온톨로지 언어에 익숙하지 않은 개발자들에게 개발을 쉽게 할 수 있도록 도와줄 것이다. 특히, 실제 상품 온톨로지 개발에 있어 우리가 제안한 새로운 모델링 요소는 매우 유용하게 사용 될 수 있을 것이다. 현재 편집도구는 프로토타입 (Prototype) 수준으로 구현되어 있는 상태이며, 앞으로는 자동 생성된 OWL 코드를 직접입력 (Editing) 할 수 있게 하여, EER ↔ OWL 간의 상호 변환을 손쉽게 할 수 있게 할 것이며, 이미 생성된 Class 레벨의 모델뿐 아니라, 인스턴스도 또한 모델링 틀에서의 직접입력이 가능하도록 발전시킬 것이다.

6. 참고문헌

1. Ig-hoon Lee, Suekyung Lee, Taehee Lee, Sang-goo Lee, Dongkyu kim, Jonghoon Chun, Hyunja Lee, Junho Shim, "Practical Issues for Building a Product Ontology System", DEEC2005, April 9, 2005.
2. Enrico Franconi and Gary Ng, "The i*com Tool for Intelligent Conceptual Modelling", KRDB, 2000.
3. <http://www.w3.org/TR/2004/REC-owl-ref-20040210>.
4. <http://protege.stanford.edu>.