

웹 온톨로지를 위한 온톨로지 구축 기법

김수경*, 안기홍*

*한밭대학교 컴퓨터공학과

e-mail:kimsk@hanbat.ac.kr

Ontology Building Methodology for Web Ontology

Su-Kyoung Kim*, Kee-Hong Ahan*

*Dept of Computer Engineering, Hanbat National University

요 약

시맨틱 웹 응용의 구현에 있어 가장 중요한 기술이 시맨틱 웹의 특징을 만족하는 웹 온톨로지의 구축임에도 불구하고, 대부분 웹 온톨로지의 구축에 적용된 기법들이 시맨틱 웹과 웹 온톨로지의 특징을 만족하지 못함에 따라 시맨틱 웹 응용의 발전과 보급이 미흡하다. 따라서 본 연구는 온톨로지와 웹 온톨로지 관련 연구들을 분석하여 시맨틱 웹을 위한 웹 온톨로지의 특징들을 파악하고, 기존 온톨로지 구축과 웹 온톨로지 구축을 위해 제안된 구축 기법들을 비교 분석하여 시맨틱 웹과 웹 온톨로지의 특징에 적합한 웹 온톨로지 구축 기법을 제안하였다.

1. 서론

시맨틱 웹 응용의 최종 단계는 현재 웹 기술을 바탕으로 하여, 웹 온톨로지의 구축을 기반으로 추론 엔진을 이용하여 지식에 대한 규칙과 추론 기능이 제공되는 시맨틱 웹의 최소 요건[1]을 만족하는 응용시스템이다. 그러나 시맨틱 웹 응용의 기반 기술인 추론 기반의 웹 온톨로지 구축은 다른 시맨틱 웹 기술 수준들에 비해 발전 수준이 부족한 상황이고 이는 시맨틱 웹 응용의 발전과 보급에 장애가 되고 있다. 또한 팀 버너스 리가 2005년 7월 AAI에서 발표[2]한 바와 같이 시맨틱 웹은 인공지능이 아님에도 불구하고, 웹 온톨로지 구축은 시맨틱 웹과 웹 온톨로지의 특징을 고려하지 않고 인공지능 분야에서 연구된 온톨로지 개념과 구축 기법이나 특정 응용을 위한 온톨로지 구축 기법 또는 비형식적이고 체계적이지 못한 온톨로지 구축 기법 등을 사용하고 있다.

따라서 본 연구는 기존 온톨로지 구축 기법들을 비교 분석하고 시맨틱 웹 특징을 파악하여 시맨틱 웹과 웹 온톨로지의 특징을 만족하는 웹 온톨로지 구축 기법을 제안하고 웹 온톨로지를 구축하여 실험하였다.

2. 웹 온톨로지 구축 기법 기초 연구

2.1 시맨틱웹에서의 웹 온톨로지

시맨틱 웹 구조에 있어서 웹 온톨로지의 특징과 역할의 조사 분석 결과는 다음과 같다. 첫째, 웹 온톨로지는 온톨로지 공학자나 기계가 이해할 수 있도록 지식이 명시적이고 형식적으로 표현되어야 한다. 둘째, 구문론적 상호작용으로 파싱된 데이터의 재사용을 증가시키거나 의미론적인 상호 작용 등을 현재 웹 환경보다 향상시켜야 한다. 셋째, 현재 웹을 바탕으로 정보의 표현과 추론이 추가되어야 한다. 넷째, 응용들 간에 웹 기반의 지식 처리와 공유, 재사용이 가능해야 한다. 다섯째, 웹 온톨로지의 구축 언어는 RDF, DAML+OIL, OWL 등과 같은 웹 온톨로지 언어를 이용한다. 여섯째, 웹 온톨로지가 존재론에서 거론하는 모든 형상에 대한 표현을 하는 것은 매우 어렵기 때문에 웹이라는 특정 분야를 표현하고 구축한다. 일곱째, 웹 온톨로지는 웹 서버를 통해 운용되고 응용시스템과 연결되며 개방형의 구조를 갖는 것이 좋다. 여덟번째, 웹 온톨로지는 온톨로지에 포함된 개념의 의미를 유추할 수 있도록 추론 엔진 시스템이 필요하다.

2.2 기존 온톨로지 구축 기법 비교 분석

인공지능 분야에서 시작된 대표적인 온톨로지 구축 기법과 시맨틱 웹의 제안된 후 발표된 온톨로지 구축 기법들에 대한 비교 분석 결과는 <표1>과 같다.

<표1> 온톨로지 구축 방법론별 비교

구분		국 외								국 내	
		Cyc	KACTUS	TOVE	Method-ontology	Ontology Development 101	OTKM	DOLCE	Casual Web Ontology	EOE	IEEB
온톨로지 생성 기법	Bottom-up	○	-	-	-	-	-	-	○	-	-
	Top-down	-	○	○	-	-	-	-	-	-	○
	middle-out	-	-	-	○	○	○	○	-	○	-
온톨로지 시각 방법	업무 중심 개발	-	○	○	-	-	-	○	-	○	-
	추상적 개발	○	-	-	○	○	○	-	○	-	○
모델	단계별 모델	○	○	○	-	○	-	○	○	-	○
	프로토타입	-	-	-	○	-	○	-	-	○	-
온톨로지 표현	형식화된 온톨로지 기술	-	-	○	○	-	○	○	○	○	-
	지식 표현 언어	F-logic, CycL	-	C++	-	DL	-	FOL	-	-	-
	지원 영역	일반	업무	업무/기법	일반	일반	프로젝트	프로젝트	일반	일반	일반
	개발년도	-	1994	1992	1997	2000	1999	2004	2004	2004	2004

각 구축 방법별 특징은 다음과 같다.

첫째, OTKM과 METHONTOLOGY와 방법은 개발 단계의 모든 과정에 대한 절차를 제시한다.

둘째, KATUS, DOLCE, Ontology Development 101 방법은 온톨로지 구축 이후의 절차를 제시하지 않았다.

셋째, Cyc, TOVE 등은 지식 기반 시스템 구축에 중점을 둔 방안으로 모든 온톨로지 구축에 적용하기는 적합하지 않다.

넷째, Ontology Development 101 온톨로지 구축 방법은 온톨로지 구축의 필요성과 경험 사례를 바탕으로 기초적인 온톨로지 구성에 대한 과정을 제공하였으며, 체계적이고 구조적인 온톨로지 구축 방법을 제공하지 않았다.

다섯째, METHONTOLOGY 기법은 반복 검토를 통한 정밀화 단계가 생략되어 있기 때문에 개발속도는 빠르지만 온톨로지의 질적 수준은 미약하다.

여섯째, OTKM 방법은 소프트웨어 공학적인 접근의 프로토타입 방법을 중심으로 대규모의 온톨로지 프로젝트를 개발하는데 적합하지만 온톨로지의 개발 시간이 많이 소모될 수 있다.

일곱째, 온톨로지 공학자나 기계가 이해할 수 있도록 형식화된 지식 표현의 제공과 추론을 통한 온톨로지 구축 방법의 제시가 부족하다.

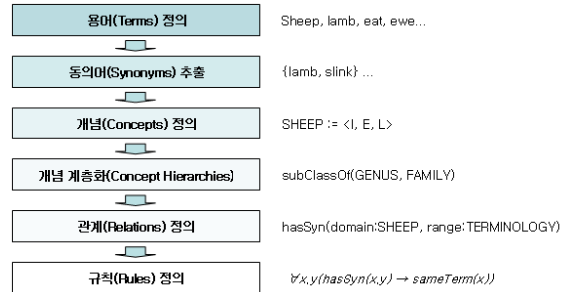
3. 웹 온톨로지 구축 기법 제안

3.1 웹 온톨로지 구축을 위한 참조 요소

(1) 지식의 명시적 표현

웹 온톨로지의 제약 사항 중 하나인 개념적 모델을 묵시적으로 표현함에 따른 코드와 표현의 차이점을 극복하기 위해 서술 논리와 규칙을 이용하여 웹

상에 존재하는 단어나 용어들로부터 규칙에 이르는 과정을 <그림1>과 같이 분석하였다. 이 단계는 제안된 웹 온톨로지 구축 기법의 온톨로지 내부 분석과 설계 단계에서 온톨로제 개념의 명시적 표현으로 활용된다.



<그림1> 온톨로지 명시적 표현 단계

<그림1>각 단계에 대한 설명은, 단계 1. 용어 계층은 온톨로지를 구축하는 지식 영역에 있는 용어들을 나열한다, 단계 2. 동의어 계층은 나열된 용어들 중에서 동의어 들을 재정의 한다. 단계 3. 개념화 단계는 동의어와 용어들을 그룹화하여 개념 또는 클래스로 표현한다. 단계 4. 개념 계층은 나열된 개념 또는 클래스들 간의 상하 관계들을 명시적으로 표현한다. 단계 5. 관계 단계는 각 개념 또는 클래스들을 도메인(domain)과 영역(range)로 결정한 뒤 개념 또는 클래스들 간의 관계를 명시적으로 표현한다. 단계 6. 규칙 단계로 이 단계에서는 단계 3과 단계 4가 서술 논리의 선언적 형식을 사용하여 지식을 표현하였다면 이 단계는 용어적 형식과 규칙 언어를 사용하여 속성과 속성간의 관계 또는 속성과 개념간의 관계 등을 명시적으로 표현한다.

(2) 시맨틱 웹 응용을 위한 온톨로지 참조 요소

<표2> 시맨틱 웹 응용을 위한 온톨로지 요소

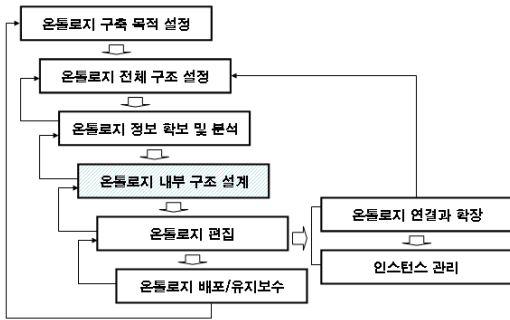
번호	주 제	설 명	확인
1	Ontology Language	구축에 사용될 온톨로지 언어	
2	Rule Language	규칙 표현에 사용될 언어	
3	Language Extension	온톨로지 확장에 필요한 언어	
4	Ontology Server Architecture	온톨로지 서버 구조	
5	Ontology Server Demonstrator	온톨로지 서버를 통한 데모	
6	Ontology Editor	온톨로지 개발 툴	
7	Ontology Inference System	온톨로지 추론 시스템	
8	Ontology DataBase	온톨로지 데이터베이스	
9	Reasoner Demonstrator	추론 데모	
10	Ontology RoadMap	구축될 온톨로지 로드맵	
11	Ontology Methodology	온톨로지 구축 기법	
12	Ontology Library	온톨로지 라이브러리 제공	
13	Module Mechanism	모듈별 동작 구조	
14	Evaluation Report	온톨로지 평가	
15	Project Presentation	온톨로지 활용 제시	
16	Technical Implementation Plan	온톨로지 구축과 응용 시스템	

<표2>의 각 주제는 시맨틱 웹 응용을 구현하였을 때 웹 온톨로지가 시맨틱 웹 응용과의 상호작용을 위해 필요한 요소이며, 이 요소들은 온톨로지를 기반

으로한 시맨틱 웹 응용의 평가 항목으로도 활용된다.

3.2 웹 온톨로지 구축 기법 제안

2장과 3.1절에 의해 조사 분석된 결과에 의해 제안된 시맨틱 웹에 적합한 웹 온톨로지 구축 기법(A Method of Building Web Ontology suitable for Semantic Web: 이하 MBWO)의 프로세스는 <그림 2>와 같다.



<그림2> MBWO 구축 기법

(1) 단계 1 : 온톨로지 구축 목적 설정

이 단계는 응용 시스템에서의 온톨로지 기능, 온톨로지 단위, 지식 표현 단위 등을 연구 조사하여 온톨로지 개요서를 작성한다. 그리고 온톨로지 구축 대상과 응용에 대한 좀 더 구체적인 정보를 기술하는 온톨로지 요약표를 작성한다. 또한 응용시스템이 얻고자 하는 질의들을 목록으로 작성한 적격 질의 목록표를 작성한다. 적격 질의 목록표는 온톨로지 내부적 구조 설계 단계에서 도메인 온톨로지의 내부를 명시적으로 구조화하고 표현하는데 사용된다.

(2) 단계 2 : 온톨로지 전체 구조 설정

이 단계는 먼저 기존에 구축된 온톨로지가 있는 경우 이를 재활용할 수 있는 방안을 검토한 뒤, 온톨로지 명세서를 작성한다. 온톨로지 명세서는 1단계에서 정의한 범위와 대상들에 대한 전체적인 온톨로지 명세와 그에 따른 각 온톨로지의 네임스페이스와 접두어(prefix) 등을 결정하며, 구축될 온톨로지들 중에서 중심 온톨로지를 결정한다. 또한 구축될 온톨로지의 하위 온톨로지들로 각 도메인 온톨로지들과 외부에서 이입될 온톨로지에 대한 정보를 정리하여 온톨로지내 각 개념들의 유일성을 보장한다.

(3) 단계 3 : 온톨로지 정보 확보 및 분석 단계

이 단계는 온톨로지 도메인 범위에 포함되는 개념들의 명세를 작성한다. 개념들의 용어와 동의어 그리고 개념에 대한 분석과 설계를 한다. 그리고 각 도메인 온톨로지별 용어(단어)를 조사한 결과인 용어조사표를 작성한다. 이 용어조사표를 기준으로 동의어나 개념 그리고 속성에 대해 개념 사전이 작성된다.

(4) 단계 4 : 온톨로지 내부 구조 설계 단계

이 단계는 개념 계층, 관계, 규칙에 대한 구체적인 정의를 위해 클래스의 속성, 제한규칙, 관계의 형태를 분석하고 도메인의 논리구조를 분석하여 규칙과 공리를 추출한 뒤 수식화 한다. 이를 위해 도메인 개념 속성표, 도메인별 개념 인스턴스표 그리고 <표3>의 도메인별 규칙 공리표를 작성한다.

<표3> 도메인별 규칙 공리표

도메인 규칙 공리표	
온톨로지명	이미지 지식 기반 온톨로지
도메인 온톨로지명	용어(terminology)
개념명	LifeStage
공리명	sameLifeStage
설명	lamb과 동일한 양의 성장 단계를 나타내는 용어를 추론한다 sameLifeStage(?x,?y) ^ sameLifeStage(?y,?z)
표현식	^ differentFrom(?x,?y) => sameLifeStage(?x,?z)
관련 개념	Sheep, Actor, LifeStage
관련 속성	sameLifeStage, differentFrom, LifeStage(x) ...

(5) 단계 5 : 온톨로지 생성과 편집 단계

이 단계에서는 온톨로지 편집 툴을 이용하여 전 단계에서 설계된 명세에 따라 실제 온톨로지를 작성한다. 도메인 온톨로지간 연결과 확장이 필요할 때는 온톨로지 연결과 확장 작업을 하고 온톨로지의 인스턴스를 생성한다.

(6) 단계 6 : 온톨로지 배포/유지 보수

마지막으로 구현이 마무리된 온톨로지는 웹상에 온톨로지의 전체 구조와 정보를 배포한다. 그리고 응용시스템의 구축, 사용자의 추가적인 요구와 도메인 지식의 갱신 등을 위해 지속적으로 이를 보완하는 유지보수가 이뤄진다.

<표4> 온톨로지 기반 응용 시스템 평가표

온톨로지 기반 응용 시스템 평가표		
번호 : _____	작성일 : _____	
응용 시스템명	지능형 이미지 검색 시스템	
응용 시스템 목적	시스템 실행 결과	비고
검색 문장을 이용해 이미지의 내용을 중심으로 이미지를 검색한다.		
적격 질의 목록	추론 결과	비고
lamb과 동일한 의미를 가진 이미지는?	slink	
merino sheep 의 종명은?	ovis aries	
.....
응용시스템과 온톨로지 요소 평가		
주 제	확 인	비고
Ontology Language	OWL, RDF	
Rule Language	SWRL, DL	
Language Extension	Java, JSP	
Ontology Server Architecture	Apach, Tomcat	
Ontology Server Demonstrator	지능형 이미지 검색	

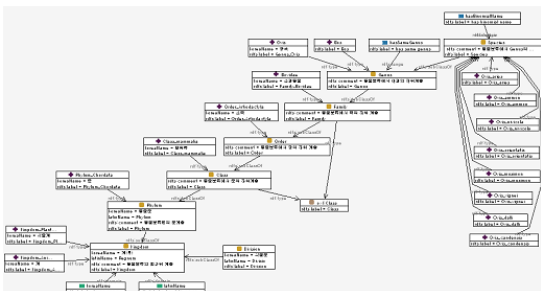
<표4>의 온톨로지 기반 응용 시스템 평가표를 작성하여 각 요소별로 평가하고 온톨로지 유지, 보수 등에 반영한다.

4. 실험

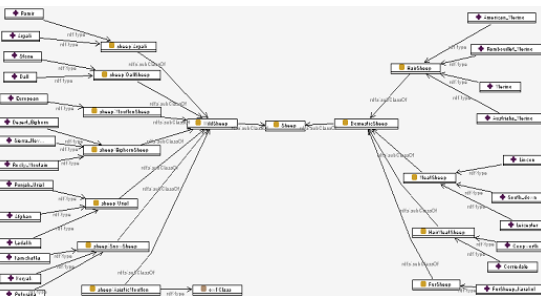
제안된 웹 온톨로지 구축 기법에 따라 구축된 도메인 웹 온톨로지를 구축하였다. 각 도메인 웹 온톨

로지들은 서로 다른 지식 영역을 표현하고 있으나 MBWO의 규칙 공리표에 의해 작성된 형식화된 개념과 속성 표현에 의해 각 도메인 온톨로지들의 지식이 이입(import)되며 이를 이용하여 서로 다른 영역의 지식이 추론(inference)된다.

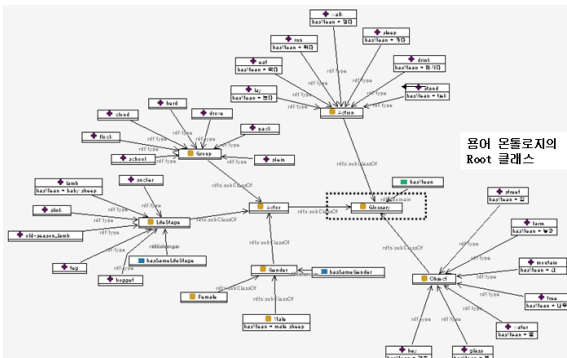
<그림3>은 동물분류 계층의 지식을 표현한 도메인 온톨로지이며, <그림4>는 동물의 한 종인 양 (sheep)에 대한 지식을 표현한 도메인 온톨로지, <그림5>는 동물이나 양에 대한 다양한 용어(terminology)를 표현한 용어온톨로지이다.



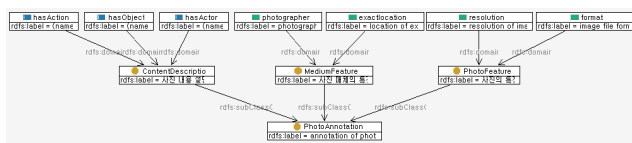
<그림3> 동물 분류 온톨로지 그래프



<그림4> 양 온톨로지 그래프



<그림5> 용어 온톨로지의 전체 구조



<그림6> 응용 프레임 온톨로지 스키마

<그림6>은 웹 온톨로지의 특징에 따라 각 도메인 온톨로지의 지식을 추론하여 응용 프레임 온톨로

지가 새롭게 구축되는 것을 보여준다. 이 응용 프레임 온톨로지는 미리 구축된 도메인 온톨로지들의 명시적이고 형식적인 지식들을 이용하여 새로운 도메인 온톨로지가 구축될 수 있는 것을 보여준다.

<그림7>은 응용 프레임 온톨로지 스키마를 통해 새롭게 완성된 온톨로지의 인스턴스 예이다..

```

:PhotoAnnotation_1
a :PhotoAnnotation ;
rdf:type :PhotoAnnotation ;
rdf:label "(name)"^^xsd:string ;
exactlocation "korea"^^xsd:string ;
format "jpeg"^^xsd:string ;
hasAction Terminology:eat ;
hasActor Terminology:lamb ;
hasObject Terminology:grass ;
resolution "100x100"^^xsd:string ;
setting "at farm"^^xsd:string .
    
```

<그림7> 생성 인스턴스의 소스 코드

5. 결론 및 향후 연구 방향

제안된 MBWO에 의해 구축된 도메인 온톨로지들은 웹 온톨로지가 갖춰야할 특징 중의 하나인 웹 온톨로지들 간의 지식 공유나 재활용과 이를 활용한 새로운 도메인 온톨로지 구축이 가능함을 보여주었다. 향후, 현재 HTML로 표현된 웹 상의 정보들을 자동으로 웹 온톨로지의 지식 표현 단위로 변환하기 위해, 제안한 MBWO 구축 기법의 단계 중 정보의 명시적 표현 단계에 의한 정보의 구조 분석 단계를 자동으로 처리할 수 있는 방안의 연구가 필요하다.

참고문헌

- [1] Michel Klein, Ubbo Visser. "Semantic Web Challenge 2003", IEEE Computer Society, IEEE INTELLIGENT SYSTEM, pp. 31-33, 2004. 5.
- [2] Tim Burners Lee. "Artificial Intelligence and the Semantic Web", <http://www.w3.org/2006/Talks/0718-aaai-tbl/>, AAAI, 2006.7.
- [3] Enterprise Integration Laboratory. "TOVE Ontology Project", <http://www.eil.utoronto.ca/enterprise-modelling/tove/index.html>, 2002.
- [4] Natalya F. Noy, Deborah L. McGuinness. "Ontology Development 101:A Guide to Creating Your First Ontology", pp. 3-9, 2001.
- [5] Dieter Fensel et al. "On-To-Knowledge: Ontology-Based Tools for Knowledge Management.", 2000.
- [6] Nicola Guarino. "DOLCE: a Descriptive Ontology for Linguistic and Cognitive Engineering", <http://www.loa-cnr.it/DOLCE.html>, 2006.
- [7] C.Masolo, S. Borp, A. Gangemi, N. Guarino, A. Oltramari. "Wonder- Web deliverable D17", intermediate report 2.0, ISTC-CNR, 2002.