

FOL 을 이용한 온톨로지 추론과 수정

강민구*, 박영택*

*송실대학교 컴퓨터학과

e-mail : happytest@korea.com, park@comp.ssu.ac.kr

Ontology Inferenceing and Updating using FOL

Min-Goo Kang*, Young-Tack Park*

*Dept. of Computer Science, Soong-Sil University

요 약

현재 웹이 가지고 있는 문제점들을 해결 하기 위해서 연구하고 있는 차세대 웹을 시멘틱 웹이라고 한다. 시멘틱 웹에서 다루고 있는 기술들은 다양하지만 시멘틱 웹 구현에 있어서 온톨로지와 그 온톨로지를 추론하여 agent 가 정보의 Semantic 을 알아내는 것이 가장 핵심 되는 영역의 기술이다. 본 논문에서는 DAML+OIL 로 작성된 온톨로지의 추론 방법과 추론 결과물을 이용한 온톨로지 수정 방법에 대해서 제안한다. 이를 위해서 온톨로지를 inference engine 에서 작업을 수행 할 수 있도록 FOL 로 변환하는 기술, DAML+OIL axiom 과 FOL 을 이용해서 inference 할 수 있는 엔진 구축 기술, inference 결과물을 DAML+OIL 로 변환하여 수정하는 방법들을 제안한다.

1. 서론

시멘틱 웹은 Agent가 웹 정보들의 semantic을 이해 하는 것을 목표로 한다. 현재 웹은 정보가 syntactic 구조 기반으로 작성되기 때문에, 사람은 그 의미를 이해 할 수 있었지만 agent는 의미를 이해 할 수 없었다. Agent란 사람이 할 수 있는 것을 대신 수행 해주는 것을 의미하는데 웹 상에서 사용되는 agent는 웹에 산재해 있는 data를 information으로 가공하는데 수많은 수작업이 뒤따르게 되어 현재 한계상태에 달한 상태이다. 이 한계상태를 극복하기 위해서 agent 가 해석할 수 있는(machine interpretable) 방식의 data 표현 방식이 필요하게 되었는데 현재 이를 위해서 DAML+OIL이 쓰이고 있다. 시멘틱 웹이 출현하게 된 중요한 동기 중의 하나는 웹을 사용하는 사용자나 웹에서 동작하는 소프트웨어 에이전트 모두가 웹 상의 정보나 서비스를 쉽게 이해하고 처리할 수 있도록 하기 위한 것이다. 에이전트가 정보와 서비스를 이해 하기 위해서는 온톨로지 구축 작업과 이 온톨로지를 추론 하여 semantic을 얻어내는 추론 기관이 필수적으로 필요하게 된다.

2. 관련연구

2.1 Stanford KSL 의 연구

Stanford KSL 연구팀에서는 DAML+OIL 온톨로지 지식을 autonomous하게 추론하기 위해서 추론 기반 질의응답기를 구축하고 있다. 소프트웨어 에이전트를 지식 서버로 보고 DAML+OIL로 표현된 지식베이스를 효과적으로 추론하여 질의 응답을 수행하기 위해서 JTP와 OQL을 활용하고 있다. 스탠포드 KSL 연구팀에서는 DAML+OIL로 표현된 정보로부터 새로운 정보를 추론하기 위해서 DAML+OIL을 FOL로 변환하는 방식을 사용한다. W3C에 제안된 DAML+OIL specification은 FOL 시멘틱스와 model-theoretic 시멘틱스를 포함하고 있기 때문에 DAML+OIL을 FOL로 변환하면 FOL 자동 추론기 (automated theorem prover)를 활용해서 추론을 수행하고 있다.

2.2 Yale대학의 연구

Yale 대 연구팀은 온톨로지 변환(translation)을 위해서 온톨로지 병합(merge)과 자동 추론 방식을 활용하는 연구를 수행하고 있다. 온톨로지 병합은 human 전문가의 도움을 받아서 이루어지고, 합병된

온톨로지의 변환은 OntoEngine 이라는 forward/backward chaining 추론기관을 활용하고 있다.

Yale 연구팀은 두개의 온톨로지의 서로 다른 term을 연관시키기 위해서 bridging axiom을 정의해서 forward chaining이나 backward chaining을 적절히 활용하는 방안을 제안한다. DAML+OIL로 표현된 내용을 FOL 형식으로 변환하고 여기에 theorem prover를 적용하는 방법이다. 이러한 방식은 기존의 인공지능 연구에서 많이 처리해온 방식으로 modus-ponens를 포함한 다양한 추론 기법을 적용하여 추론을 진행한다.

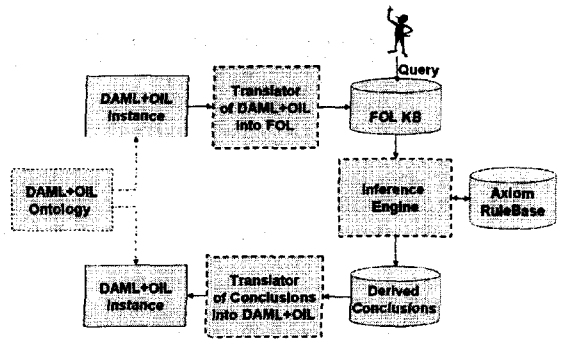


그림 1 시스템구조도

2.3 MIT 의 연구

MIT는 DAML+OIL기반의 rule표현 언어인 DamiRuleML을 이용하여 추론을 수행한다. DamiRuleML 은 시멘틱웹이 지향하고 있는 자동 추론 시스템을 개발하기 위해서는 꼭 필요한 기술이지만 시멘틱웹 연구는 온톨로지 마크업 언어에 치중한 나머지 rule에 대한 연구는 부족한 실정이었다. MIT에서의 DamiRuleML 연구는 시멘틱웹 연구에 있어서 온톨로지 마크업 언어와 rule언어 사이의 차이를 줄이기 위한 목표를 가지고 발족되어 온톨로지에 있어서 rule을 표현하기 위한 방법들을 연구한다.

DamiRuleML은 이기종 시스템 사이의 rule 교환을 가능하게 하고 rule 지식 표현을 위해 XML 문법을 사용하므로 상호 운용성이 풍부한 장점이 있다. 현재 MIT에서는 DamiRuleML을 기존의 추론엔진인 Jess에 알맞은 형태로 변환한 후 추론작업을 수행하는 SweetJess(Semantic WEb Enabling Technology)프로젝트를 진행하고 있다.

2.4 휴렛팩커드의 연구

휴렛팩커드는 Jena 라는 이름의 시멘틱 웹 종합 툴킷을 개발하고 있다. Jena 에는 RDQL 이라는 RDF 질의 인터페이스, DAML+OL 의 온톨로지를 생성하고 간단한 inference 를 제공하는 API 를 포함하고 있다.

3 온톨로지 추론 방법 연구

3.1 시스템 구조

본 시스템에는 FOL 추론을 위해서 온톨로지를 FOL 형식으로 바꾸어주는 translator와 FOL과 axiom rulebase를 이용해서 추론을 수행하는 추론 엔진, 추론 엔진을 통해서 도출된 결과를 다시 온톨로지에 첨가해 주는 translator가 시스템을 구성하는 주요 컴포넌트들로 되어 있다. 본 시스템에서는 온톨로지를 작성 하기 위해서 DAML+OIL의 기술을 사용하였으며, FOL 표현을 위하여 KIF를 이용한다.

3.2 DAML+OIL Knowledge Base

DAML+OIL은 W3C에서 권장하는 표준 온톨로지 작성 언어이다. DAML+OIL을 이용하여 온톨로지 KnowledgeBase를 작성 할 수 있다. 아래의 예는 꽃에 대한 온톨로지의 일부분으로써 장미에 대한 class와 property들이 정의 되어 있다.

```

<daml:Class rdf:ID = "Rose">
  <rdfs:subClassOf rdf:resource = "Flower"/>
  <rdfs:subClassOf>
    <daml:Restriction>
      <daml:onProperty rdf:resource = "hasRoseColor"/>
      <daml:toClass rdf:resource = "RoseColor"/>
    </daml:Restriction>
  </rdfs:subClassOf>
</daml:Class>

<Rose rdf:ID = "MyFavoriteFlower">
  <hasRoseColor rdf:resource = "Red"/>
</Rose>
  
```

그림2 DAML+OIL 온톨로지

3.2 FOL 변환

본 연구에서는 추론을 수행하기 위하여 FOL(First Order Logic)을 이용한 추론을 이용한다. 추론에서 FOL을 이용하는 이유는 기존의 theorem prover와 problem solver를 사용하기 쉽고, DAML+OIL의 확장된 문법이 새롭게 정의 되었을 때 유연하게 적용 할 수 있기 때문이다. 본 연구에서는 DAML+OIL을 FOL로 변환하기 위해서 DAML+OIL을 RDF statement로 변환하고 RDF statement를 FOL로 변환하는 simple한 rule을 이용한다. 변환 할 때 RDF statement를 이용하는 이유는 DAML+OIL, RDFS들이 모두 RDF statement의 적용을 받는 서브셋이기 때문이다.

● 변환순서

1. DAML+OIL로 표현된 온톨로지 Knowledge Base를 RDF statement로 변환한다.

```
(rdf:type Rose daml:Class)
(rdfs:subClassOf Rose Flower)
(rdf:subClassOf Rose GnR)
(rdf:type GnR daml:Restriction)
(dam1:onProperty GnR hasRoseColor)
(dam1:toClass GnR RoseColor)

(rdf:type MyFavoriteFlower Rose)
(hasRoseColor MyFavoriteFlower Red)
```

그림3 온톨로지의 RDFstatement변환

2. 변환된 RDF statement를 property P, subject S, object O로 parsing한 다음, (PropertyValue P S O) 형태로 재 변환한다. 이때 PropertyValue에는 property와 entity에 관한 relation이 정의 되어 표현한다.

```
(Type Rose daml:Class)
(PropertyValue rdfs:subClassOf Rose Flower)
(Type GnR daml:Restriction)
(PropertyValue rdfs:subClassOf Rose GnR)
(PropertyValue dam1:onProperty GnR hasRoseColor)
(PropertyValue dam1:toClass GnR RoseColor)

(Type MyFavoriteFlower Rose)
(PropertyValue hasRoseColor MyFavoriteFlower Red)
```

그림4 온톨로지의 FOL변환

3.3 Axiom Rule Base

DAML+ OIL axiom은 ANSI Knowledge Interchange Format(KIF)를 이용해서 표현한다.

```
(<=>(Type ?r ?v)(PropertyValue type ?r ?v))
```

그림 5 KIF의 axiom 표현방법

DAML+ OIL이나 RDFS는 RDF를 확장한 기술이므로 RDF의 axiom을 기반으로 모든 표현이 가능하다. Axiom은 RDF의 class와 property들을 정확하게 정의해 준다.

```
(<=>(PropertyValue subClassOf ?csub ?csuper)
  (and(Type ?csub rdfs:Class)
    (Type ?csuper rdfs:Class)
    (forall(?x)(=>(Type ?x ?csub)
      (Type ?x ?csuper))
```

그림 6 subClass axiom

3.4 Inference

본 연구에서는 온톨로지 추론을 수행하기 위해서 FOL로 변환된 DAML+OIL KnowledgeBase와 axiom RuleBase를 이용하는 forward chain 추론 엔진을 사용한다. 본 시스템에서 (Type Red ?Type)라는 질의는 아래와 같이 추론된다.

그림 4 의 FOL 로 변환된 KnowledgeBase 를 보면 Red 의 type 이 RoseColor 라는 predicate (Type Red RoseColor)는 assert 가 안되어 있으므로 (Type Red ?Type)이라는 질의어를 처리하는 것이 불가능하다. 본 시스템에서는 추론을 위해서 DAML+OIL axiom 을 forward chaining 추론함으로써 (Type Red RoseColor)를 추론하게 된다. 현재 DAML+OIL 은 10 개 class 의 14 개의 axiom, 26 개 property 의 69 개의 axiom 이 정의되어 있다.

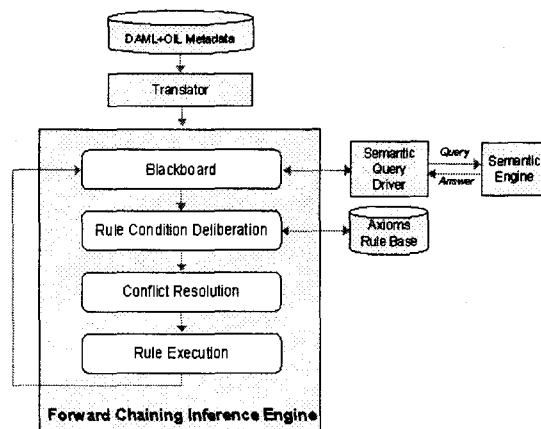


그림 7 Inference Engine

본 시스템의 Inference Engine 은 쿼리를 수행 하기 위해서 subClassOf axiom 을 먼저 사용한다.

```
(<=>(PropertyValue subClassOf ?csub ?csuper)
  (and(Type ?csub rdfs:Class)
    (Type ?csuper rdfs:Class)
    (forall(?x)(=>(Type ?x ?csub)
      (Type ?x ?csuper))
```

Inference Engine 은 위의 axiom 과 앞에서 변환된 FOL facts 들에 unification 을 적용하면서

Rose 를 ?csub 로 치환(substitution)하고, GnR 을 ?csuper 로 치환하여 다음과 같은 결과를 유도한다.

```
(=>(PropertyValue subClassOf Rose GnR))
  (=(>(Type MyFavoriteFlower Rose)
      (Type MyFavoriteFlower GnR)))
```

(PropertyValue subClassOf Rose GnR) 과 (Type MyFavoriteFlower Rose) 이 위의 mapping 된 statement 에서 이미 존재 하므로 reasoner 는 (Type MyFavoriteFlower GnR)을 reasoning 할 수 있게 된다. 이 결과와 아래의 toClass axiom 을 이용하면 Red 가 RoseColor type 임을 reasoning 할 수 있게 된다.

```
(=>(and(PropertyValue onProperty ?r ?p)
        (PropertyValue toClass ?r ?c))
    (forall(?i)
      (<=>(Type ?i ?r)
        (forall(?i)
          (>(PropertyValue ?p ?i ?j)
            (Type ?j ?c))))))
```

Reasoner 가 GnR 을 ?r 으로, hasRoseColor 를 ?p 로, RoseColor 를 ?c 로, MyFavoriteDrink 를 ?i 로 치환하면 toClass axiom 은 아래와 같은 결과를 도출하게 된다.

```
(=>(PropertyValue hasRoseColor MyFavoriteFlower ?j)
    (Type ?j RoseColor))
```

따라서, Reasoner 가 Red 를 ?j 와 치환하게 되면 Inference Engine 에서 query 로 입력된 (Type Red RoseColor)를 추론하여 결과를 output 하게 된다.

3.5 DAML+OIL 수정

추론으로 derive된 결과는 DAML+OIL 형태로 변환되어 DAML+OIL KnowledgeBase에 추가된다. 변환되는 과정은 추론하기 위해서 FOL 형태로 변환한 과정의 역순으로 일어 나게 되는데 위의 추론으로 인하여 도출된 (Type Red RoseColor)는 <RoseColor rdf:ID = "Red"/>의 DAML+OIL 형태로 변환되어 KnowledgeBase에 추가된다. 수정된 DAML+OIL 온톨로지는 전과 같은 query가 들어 왔을 때 자원이 소비되는 inference 과정을 거치지 않고 결과를 도출 할 수 있게 된다.

4 결론 및 향후 과제

본 연구에서는 온톨로지 추론방법과 추론결과물을 이용한 온톨로지 수정에 대한 방법을 제안하였다.

온톨로지가 아무리 견고하게 작성 되었다고 하더라도 온톨로지를 추론하는 장치가 없다면 온톨로지는 제 역할을 할 수 없다. 본 시스템은 온톨로지 추론에 대한 기본적인 방법을 제안하는데 그 의의를 갖는다. 하지만 FOL을 이용한 reasoning은 maxCardinality와 같은 cardinality restriction 추론하는데 많은 문제점을 가지고 있고, 현재 W3C에서 표준안으로 연구하고 있는 rule 기능이 부여된 기술이나 OWL같은 새로운 기술에 의해 표현되는 온톨로지를 추론하지 못하고 있다. 향후 연구는 이점에 착안하여 이루어 져야 할 것이다.

참고문헌

- [1] Benjamin N. Grosf, Mahesh D. Gandhe, and Timothy W. Finin, " SweetJess: Translating DamLRuleML to Jess" , MIT Sloan School of management, 2002.
- [2] Benjamin N. Grosf, "Representing E-Business Rules for the Semantic Web : Situated Courteous Logic Programs in RuleML", MIT Sloan School of Management, Dec., 2002.
- [3] Benjamin N. Grosf, "Standardizing XML Rules : Preliminary Outline of Invited Talk", MIT Sloan School of Management, 2000.
- [4] Deborah L. McGuinness, Richard Fikes, James Hendler, and Lynn Andrea Stein, "DAML+OIL:An Ontology Language for the Semantic Web", IEEE Intelligent Systems, September, 2002.
- [5] F. van Harmelen et al., "A Model-Theoretic Semantics for DAML+OIL(March 2001)", W3C Note 18, World Wide Web Consortium, Dec. 2001; www.w3.org/TR/2001/NOTE-daml+oil-model20011218
- [6] I. Horrocks and U. Sattler, and S. Tobies, "Practical reasoning for expressive description logics", number1750 in LNAI, pp. 161-180, Springer-verlag, 1999.
- [7] Richard Fikes, and Deborah L McGuinness, "An Axiomatic Semantics for RDF, RDF-S, and DAML+OIL", Stanford University, March, 2001.
- [8] D. Connolly et al., eds, Annotated DAML Ontology Markup, World Wide Web Consortium, Note 18, Dec. 2001, www.w3.org/TR/daml+oil-walkthru
- [9] Richard Fikes, Deborah McGuinness, and Sheila McIlraith, "Tools for DAML-Based Services, Document Templates, and Query-Answering", Stanford University, 2002.
- [10] Sean Bechhofer, Ian Horrocks, Carole Goble, and Robert Stevens, "OilEd " a Reason-able Ontology Editor for the Semantic Web", University of Manchester, 2001.