

테이블로 알고리즘 기반 OWL 온톨로지 추론 엔진

김제민^o 권순현 최정화 박영택

송실대학교 컴퓨터학과

kimjemins@hotmail.com, kwonshzzang@naver.com, cjh79@ailab.ac.kr, park@ssu.ac.kr

Tableaux Algorithm based OWL Ontology Reasoner

Je-Min Kim, Sun-Heon Kwon, Jung-Hwa Choi, Young-Tack Park

School of Computing, Soongsil University

온톨로지가 대용량화됨에 따라, 구축 과정에 많은 인력이 투입되고, 그 과정 역시 복잡해지고 있다. 따라서 온톨로지 구축과정에서 발생하는 여러 가지 논리적 오류를 찾아내어 수정하는 작업은 중요하다. 또한 온톨로지 기반의 검색이나 온톨로지들을 통합할 때 온톨로지를 구성하는 개념간의 관계를 추론하는 것 역시 매우 중요하다. 본 연구의 목표는 온톨로지 구축 시 논리적 오류를 갖는 개념들을 찾아주고, 개념들 간에 관계를 추론하는 엔진을 구축하는 것이다. 본 논문에서 제안하는 Minerva는 OWL로 작성한 온톨로지 중 논리적 오류를 갖는 개념들을 찾아내어, 온톨로지 개발자들이 효과적으로 온톨로지를 구축하는 것과, 개념간의 관계를 추론해 줌으로써 온톨로지 기반의 서비스 어플리케이션 구축에 도움을 준다.

1. 서론

차세대 웹 환경인 시맨틱 웹에서 중추적인 역할은 하는 것은 기계가 이해할 수 있도록 개념들을 명확하게 명시하고, 이들 개념들을 공유할 수 있는 형식으로 표현한 온톨로지이다. 온톨로지가 대용량화됨에 따라, 구축 과정에 많은 인력이 투입되고, 구축 과정 역시 복잡해지고 있다. 따라서 온톨로지 구축 작업 중에 수많은 논리적 오류들이 발생되며, 이러한 논리적 오류들을 수정하는 것은 매우 중요한 작업이 되었다. 뿐만 아니라 온톨로지 기반의 시맨틱 검색이나 온톨로지들을 통합 할 때, 온톨로지들을 구성하는 개념들 간의 관계를 추론하는 것 역시 매우 중요하다. 현재 온톨로지의 논리적 오류와 개념들 간의 관계를 도출하기 위해서, OWL 추론 엔진들이 소개되고 있다[1][2][3]. 현재 소개되고 있는 대부분의 추론엔진들은 Tableaux 알고리즘[4]을 기반으로 구축되어져 있다. Tableaux 알고리즘기반의 온톨로지 추론 엔진은 규칙 기반 온톨로지 추론 엔진에 비해서 Sound하면서 Complete한 결과를 보여준다. 현재 국내에서는 규칙기반의 온톨로지추론 엔진인 보쌈[8]을 제외하고 Tractable하게 수행할 수 있는 온톨로지 추론 엔진이 아직 개발되지 않고 있다.

본 연구의 목표는 온톨로지의 논리적 오류를 유발하는 개념들을 찾아주고, 온톨로지를 구성하는 개념들 간의 관계를 Sound하면서 Complete하게 추론하는 온톨로지 엔진인 미니바(Minerva)를 구축하는 것이다. 본 논문에서 제안하는 Minerva는 OWL로 작성한 온톨로지 중 논리적 오류를 갖는 개념들을 찾아내어, 온톨로지 개발자들이 효과적으로 온톨로지를 구축하는 것과, 개념간의 관계를 추론해 줌으로써 온톨로지 기반의 서비스 어플리케이션 구축에 도움을 준다.

2. 본론

Minerva는 서술 논리(Description Logic) 기반의 OWL 온톨로지 추론엔진이며, 현재 SHI 수준의 표현력을 지닌 온톨로지의 T-Box 추론 기능 지원한다. 현재 Minerva는 온톨로지의 논리적 오류가 존재하는지를 검사하고, 포함 관계 추론을 통해 온톨로지를 구성하는 개념들 간의 계층 구조를 계산한다.

• 온톨로지의 논리적 오류 검사 과정 - 온톨로지의 논리적 오류 검사는 T-Box에 저장된 개념들을 하나씩 검증하는 것이다. 각 개념을 검증할 때 Tableaux 알고리즘을 진행시키면서, 해당 규칙을 하나씩 적용한다. 이때 규칙을 적용하는 순서가 매우 중요하다. Minerva는 스케줄링 큐인 To-do 리스트 적용하여 적용되는 규칙의 순서를 지정

한다. To-do 리스트는 규칙이 적용될 개념과 그 개념이 위치한 노드를 기록하고 있는 큐들의 집합으로서, 각각의 큐는 서로 우선순위를 가지고 있다. 따라서 Tableau 알고리즘을 진행되면서 노드개수가 확장 될 경우, 규칙을 적용할 개념들이 노드 레이블에 존재하는지 모든 노드를 검사할 필요가 없다. 검증하고자하는 각 개념에 대해 Tableau 알고리즘을 진행하면서, 논리적 모순(Contradiction)이 발생(같은 노드 레이블에 개념C와 부정인 $\neg C$ 가 동시에 존재)하면 논리적 오류 개념을 캐싱(6장에서 설명)하고, 논리적 오류를 일으키는 개념으로 분류하여, 인터페이스를 통해 온톨로지 구축 자에게 보이게 된다. 이때 검증을 마친 개념들에 대해서 온톨로지 추론 과정을 보다 빨리 수행하기 위한 의사 모델(Pseudo Model)을 구축하게 되는데, 의사 모델은 Tableau 알고리즘이 진행되면서 생성된 트리의 루트 노드 레이블 정보로 구성된다.

▪ 온톨로지 추론 과정 - Minerva에서의 온톨로지 추론은 두 개념간의 목시적 포함관계를 추론하는 것이다. 일반적으로 포함관계 추론을 시도하는 횟수는 n 개 개념이 존재할 때 $(n^2-n)!$ 번이다. 포함관계 검증은 많은 시간을 소모하므로, 온톨로지에 명시된 계층 구조를 이용하여 미리 순서리스트를 만들어 두면 포함관계 검증 횟수를 줄일 수 있다. Minerva에서 순서리스트 역할을 해주는 것이 정의 순서(Definition Order)인데, 정의 순서는 온톨로지 상의 두 개념 x 와 y 가 $x \sqsubseteq y$ 의 관계를 가질 때, $x < y$ 의 순서로 개념들을 나열해준다. 정의 순서를 이용한 포함관계 검증 횟수는 $n(n+1)/2$ 번이다. 정의 순서가 정해지면 계층 구조를 구축하며 개념들 간 관계를 탐색하는 알고리즘을 이용하여 목시적인 포함 관계를 검증하고, 이를 바탕으로 새로운 계층 구조를 구축한다. 두 개념의 포함관계 $C \sqsubseteq D$ 를 증명하기 위해서는, 이것의 논리 부정인 $C \sqcap \neg D$ 가 정당하지 못하다는 것을 보이면 된다. 일반적으로 Tableau 알고리즘은 많은 시간과 공간을 소모한다. 따라서 $C \sqcap \neg D$ 에 대해서 Tableau 알고리즘을 적용하여 정당성을 테스트하기에 앞서, 논리적 오류 검사 과정에서 생성한 C 와 $\neg D$ 개념의 의사 모델을 하나의 노드로 합병하여, 논리적 모순발생 하는지 확인한다. 만약 논리적 모순이 발생한다면, Tableau 알고리즘적용 없이 논리적으로 정당하지 않다고 판단할 수 있다.

3. 결론

현재 국내에서는 Sound하면서 Complete한 추론 결과를 보이는 온톨로지 추론 엔진이 개발되고 있지 않다. 본 연구의 목표는 온톨로지의 논리적 오류를 유발하는 개념들을 찾아주고, 온톨로지를 구성하는 개념들 간의 관계를 Sound하면서 Complete하게 추론하는 온톨로지 엔진인 미너바(Minerva)를 구축하는 것이다. 미너바가 효율적으로 동작하게 하기위해 이미 발표되었던 최적화 기법 외에 논리적 오류 개념 캐싱과 개선된 상-하위 검색, 하위-상 알고리즘을 적용하였다. 실험 결과 미너바는 이미 뛰어난 성능을 증명했었던 pellet과 큰 차이 없는 성능을 보여주고 있다. 현재 미너바는 Non-deterministic한 부분 처리에 있어서 Dependency Direct Backtracking(이하 DDB) 대신에 Chronological Backtracking 방식을 사용하고 있다. 따라서 향후 미너바는 ATMS를 적용하여 기존 추론엔진들의 DDB방식보다 효과적인 DDB 기능과 논리적 오류가 발생할 경우, 자동으로 원인을 찾아내고 디버깅하는 기능이 연구 적용될 것이다.

[참고문헌]

- [1] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur and Yarden Katz. Pellet: A practical OWL-DL reasoner, Journal of Web Semantics, 5(2), 2007.
- [2] Dmitry Tsarkov and Ian Horrocks, FaCT++ description logic reasoner: System description, In Proc. of the Int. Joint Conf. on Automated Reasoning, IJCAR, 2006.
- [3] V. Haarslev and R. Möller. Racer: A Core Inference Engine for the Semantic Web. In Proceedings of the 2nd International Workshop on Evaluation of Ontology-based Tools (EON2003), located at the 2nd International Semantic Web Conference ISWC 2003, Sanibel Island, Florida, USA, October 20, pages 27-36, 2003.
- [4] Franz Baader and Ulrike Sattler, An Overview of Tableau Algorithms for Description Logics, Studia Logica, Volume 69, Number 1.
- [5] Ian Horrocks, Optimizing Tableau Decision Procedures for Description Logics, PhD thesis, University of Manchester, 1997
- [6] Franz Baader, Bernhard Hollunder, Bernhard Nebel, Hans-Jürgen Profitlich, An Empirical Analysis of Optimization Techniques for Terminological Representation System, Principles of Knowledge Representation and Reasoning - Proceedings of the 3th International Conference, October 1992, Cambridge, MA.
- [7] U. Hustadt, B. Motik U. Sattler. Reasoning in Description Logics with a Concrete Domain in the Framework of Resolution. Proc. of the 16th European Conference on Artificial Intelligence (ECAI 2004), August, 2004, Valencia, Spain, pp. 353-357.
- [8] Minsu Jang, Joo-chan Sohn, Bossam: an extended rule engine for OWL Inferencing, Proceedings of RuleML 2004 (LNCS Vol.3323), Nov. 8, 2004