

온톨로지 포함관계 추론을 위한 최적화된 검색방법

권순현[○] 김제민[○] 박영택

송실대학교

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

An Optimized Searching Method for Ontology Subsumption Reasoning

SoonHyon Kwon[○] Jemin Kim[○] YoungTack Park

Soongsil University

1. 서 론

온톨로지의 스키마 레벨 추론에서 중요한 능력이 자동분류(Classification)이다[1]. 자동분류법은 개념들간의 계층구조를 자동으로 계산해주는 추론능력이다. 즉 개념 정의에 기반한 개념들간 상위·하위 개념 계층구조를 자동으로 만들어준다. 이러한 추론된 계층구조는 정적인 지식을 인텔리전트한 지식으로 확장해주는 역할을 한다. 본 논문에서는 온톨로지 스키마 레벨 추론의 중요한 부분인 자동분류법 추론의 최적화 기법을 제시한다.

2. 본 론

온톨로지 자동정렬법의 부분정렬은 반복적으로 만들어진다. 이러한 정렬은 초기에 $\perp < \top$ 로 생성된다[2]. 새로운 개념 C가 개념 계층구조(concept hierarchy)에 추가 되었을 때 모든 과정은 반복적으로 진행된다. 추가된 각각의 개념은 자기 자신의 직접적인 상위개념(predecessor)의 집합과 직접적인 하위개념(successor)의 집합을 결정한다. 이러한 두 개의 개념집합으로부터 추가된 개념C의 개념 계층상위 위치를 결정한다. 상위개념의 집합은 하향식 검색방법(top-down search)으로 진행되어지고 반대로 하위개념의 집합은 상향식 검색방법(bottom-up search)로 진행되어진다.

앞에서도 언급했듯이 포함관계 테스트는 복잡도가 높다. 자동분류법 과정의 전체적인 성능은 가장 적은 횟수의 포함관계 테스트를 수행하는 것이다. 본 논문의 목적은 하향식·상향식 검색방법을 최적화하여 포함관계 테스트의 횟수를 줄임으로써 자동분류법의 성능을 향상시키는 것이다. 이것을 위해 다음과 같은 기본적인 아이디어를 제공한다.

- 부정정보 하향전파(negative information down propagation) : 개념 A가 분류되었을 때 하향식 검색에서 부정정보(negative information)는 개념 계층의 하위방향으로 전파된다.
- 긍정정보 상향전파(positive information up propagation) : 개념 A가 분류되었을 때 하향식 검색에서 긍정정보(positive information)는 개념 계층의 상위방향으로 전파된다.
- 부정정보 상향전파(negative information up propagation) : 개념 A가 분류되었을 때 상향식 검색에서 부정정보(negative information)는 개념 계층의 상위방향으로 전파된다.
- 긍정정보 하향전파(positive information down propagation) : 개념 A가 분류되었을 때 상향식 검색에서 긍정정보(positive information)는 개념 계층의 하위방향으로 전파된다.
- 하향식 검색에서 긍정정보 추가(positive information gain for top-down search) : 개념 A가 개념 B에 포함된다면 개념 B를 포함하는 모든 개념 역시 A를 포함한다.
- 하향식 검색에서 부정정보 추가(negative information gain for top-down search) : 개념 A가 개념 B에 포함되지 않는다면 개념 B에 포함되어지는 모든 개념 역시 개념 A를 포함하지 않는다.

- 상향식 검색에서 긍정정보 추가(positive information gain for bottom-up search) : 개념 A가 개념 B를 포함한다면 개념 B에 포함되는 모든 개념들도 개념 A를 포함한다.
- 상향식 검색에서 부정정보 추가(negative information gain for bottom-up search) : 개념 A가 개념 B를 포함하지 않는다면 B에 포함되는 모든 개념들로 개념 A를 포함하지 않는다.

```

enhanced-top-subst?(y,c) =
  if marked(y, "positive") then
    Result ← true
  elseif marked(y, "negative") then
    Result ← false
  elseif
    exists z with z <i y
      marked?(z,"positive") then
        mark(y,"positive")
        Result ← true
  elseif for all z with y <i z
    enhanced-top-subst?(z,c)
    and subst?(y,c)
    then
      mark(y, "positive")
      Result ← true
  else
    mark(y, "negative")
    Result ← false
  fi
  fi
  
```

그림 1. 최적화된 하향식 검색 알고리즘

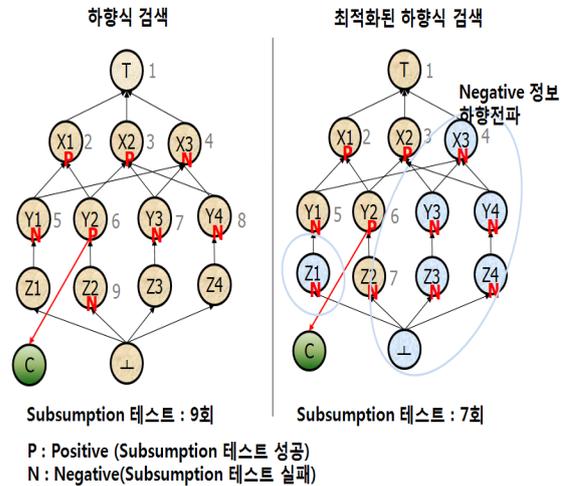


그림2. 하향식 검색과 최적화된 하향식의 비교

그림1은 위에서 제시한 기본적인 아이디어를 바탕으로 구현된 최적화된 하향식 검색 알고리즘을 나타내고 그림2은 하향식 검색방법과 최적화된 하향식 검색방법을 비교한 그림이다그림2에서 최적화된 하향식 검색방법이 2회 포함관계 테스트 횟수가 감소된 것을 볼 수 있다

아래의 표1은 Minerva & Pellet의 성능(classification, consistency)를 비교한 것이다 위의 실험은 CPU (Intel(R) Core2 CPU 660@ 2.40GHZ) 와 메모리(2G)의 PC에서 실험한 결과이다

온톨로지	개념#	최적화검색	일반검색
koala.owl	21	124	131
Pizza.owl	76	782	1,220
miniEconomy.owl	337	15,931	20,978
miniTransport.owl	444	31,967	33,386

표 1. 하향식 검색에 대한 성능평가표

3. 결론

우리는 온톨로지에 속한 개념에 대한 계층구조를 자동으로 계산해 주기 위한 자동분류법의 최적화 기법에 대해 논했다. 이를 위해 최적화된 하향식 검색 알고리즘을 적용하였고 성능평가를 통해 입증하였다. 결론적으로 최적화 기법은 복잡하고 비용이 큰 포함관계 테스트의 횟수를 줄임으로써 진행된다

현재 온톨로지 기반 추론의 복잡도 및 정해진 시간내에 결과를 도출하지 못하는 문제에 대한 최적화 기법에 대한 연구가 많이 진행되고 있다 특히 스키마 수준의 자동분류법 추론에 대한 최적화 기법은 전체적인 추론성능의 향상에 결정적이라고 할 수 있다

참고 문헌

[1] Daniele Nardi, Ronald J. Brachman, "An Introduction to Description Logics" description Logics Handbook 2003, pp. 5~6.
 [2] Frantz Baader, Brenhard Hollunder, "An Empirical Analysis of Optimization Techniques for Terminological Representation Systems". October 1992