

DBMS 기반의 온톨로지 확장 모델

이미경^o 김평 정한민 성원경
한국과학기술정보연구원 정보시스템연구팀
{jerryis^o, pyung, jhm, yksung}@kisti.re.kr

Extended Ontology Model based on DBMS

Mikyong Lee^o, Pyung Kim, Hanmin Jung, Won-Kyung Sung
Information System Research Team, Korea Institute of Science & Technology Information

요 약

본 논문은 시맨틱 웹 기술이 융합된 지식기반 정보유통 플랫폼(OntoFrame-K[®])의 추론 서비스 시스템(OntoThink-K[®])에서 이용되는 Persistent Model인 DBMS기반의 온톨로지 확장 모델에 대해 설명하고자 한다. OntoFrame-K[®]는 대용량의 지식 데이터를 다루기 때문에 기존에 개발된 온톨로지 추론 엔진을 이용할 경우 많은 한계점을 가지게 된다. 따라서 우리는 대용량의 지식 데이터를 안정적으로 처리할 수 있으며 추론의 신뢰성과 정합성을 가지는 온톨로지 확장 모델을 설계, 구현하였다. 본 모델은 OWL과 인스턴스 데이터를 트리플 형태로 변환하여 입력 받은 후, 온톨로지 스키마 규칙과 사용자 정의 규칙을 이용한 전방향 추론 방법으로 추론 서비스에서 필요한 지식데이터들을 생성하는 역할을 한다. 본 모델은 DBMS를 이용하여 대용량의 지식 데이터를 저장할 수 있으며, 추론 규칙에 따른 전방향 추론을 통해 지식 모델을 확장하기 때문에 데이터의 정합성이 보장된다.

1. 서 론

과학 기술이 발달함에 따라 같은 분야 연구자들간의 협력 연구를 지원할 수 있도록 정보 자원 공유에 기반을 둔 협력 연구 체제를 구축해야 할 필요성이 커지고 있다. 과학기술 경쟁력 강화와 연구 개발 활동의 생산성 향상을 위해서는 기존의 정보획득 과정의 개선이 필요하다. 유통 대상 정보 자원이 증가하면서 정보의 형태적 유사성에 기초한 기존의 정보 유통 체제의 기술적 한계가 드러남에 따라, 이를 해결하기 위해서 대량의 정보 자원으로부터 수요자가 원하는 정보만을 선별하여 제공할 수 있는 의미 기반의 정보 유통 기술이 대두되고 있다. 이에 적합한 대안으로 기존의 정보 유통 체제에 시맨틱 웹과 온톨로지를 접목하는 기술이 나타나고 있다[1].

이와 같은 변화에 대응하기 위해 한국과학기술정보연구원(KISTI)은 2004년부터 지식기반 정보유통 플랫폼(OntoFrame-K[®])을 개발하여 정보 유통 기술과 시맨틱 웹 기술이 융합된 새로운 개념의 정보유통 플랫폼을 개발하고 있다. 이 플랫폼은 의미 기반의 검색과 추론이 가능하고 서로 이질적인 지식 정보들을 온톨로지로 구축함으로써 지식의 재사용성을 높이고 지식 정보의 공유 효과를 누리게 된다. 그리고 온톨로지 기반의 추론 서비스를 이용하여 연구자 간 협업에 필요한 정보와 서비스를 제공하는 역할을 한다[2].

본 논문에서는 OntoFrame-K[®]의 추론 서비스 시스템(OntoThink-K[®])에서 사용되는 DBMS기반의 온톨로지 확장 모델에 대해 설명하고자 한다.

2. 관련 연구

2.1 Sesame

RDF 저장소인 Sesame는 RDBMS 기반의 RDFS 추론 엔진을 제공한다[3]. 이 시스템은 데이터 저장 과정에서 RDFS의 규칙에 따른 추론 과정을 거쳐서 결과 값을 데이터베이스에 저장해둔다. 이 시스템은 전방향 추론 방법을 이용하기 때문에 질의 처리 속도가 빠르고 질의 처리 구현 자체가 매우 간단해진다. 하지만 데이터를 저장하는 단계에서 추론을 수행하기 때문에 데이터 저장 속도가 오래 걸리는 단점이 있다. 그리고 RDFS 단계의 추론을 제공하기 때문에 OWL의 경우 공리들에 대한 표현력을 지원하지 않으며, 별도의 규칙 엔진을 제공하지 않기 때문에 사용자가 정의한 규칙은 처리할 수 없는 단점을 가진다.

Sesame는 RDFS entailment rules를 이용한 전방향 추론을 수행하기 때문에 규칙들간의 의존 관계로 인해 추론 과정에서 중복된 결과 값이 생성되는 단점이 있다. 이 점은 데이터 저장 성능의 확장성에 관한 문제점으로 대두되고 있다. 이와 관련된 Sesame의 추론 성능을 개선하기 위해

RDFS entailment rules 적용 순서를 이용한 RDFS 추론 엔진의 최적화 연구가 진행되고 있다[4].

2.2 Jena

Jena는 HP 연구소에서 만든 시맨틱 웹 프레임워크로 규칙 기반 추론 엔진을 포함하여 RDF, RDFS 및 OWL 등 온톨로지 관련 언어를 지원하는 시맨틱 웹 추론 기술을 활용한 시스템 구현에 적당한 환경을 제공한다[5]. Jena는 RDF 형태에서 시작하여 OWL과 같은 온톨로지 언어에 대한 지원을 통해 그 기능이 확장되었다. RDF/OWL API에는 간단한 메소드 호출을 통해 RDF 모델을 RDF/XML이나 N3, N-Triple 형태로 읽어들이거나 활용할 수 있는 방법을 제공한다. Jena는 메모리 기반의 추론과 RDBMS 기반의 추론을 제공하고 있다. DBMS 기반의 Jena는 Triple들의 Persistent Model을 생성하여 RDBMS에 저장한다. 그 이후, 추론 엔진은 메모리상의 모델이 아닌 Persistent Model에 접근하여 추론 결과를 도출함으로써 대용량 온톨로지의 경우 잦은 로딩 과정을 거치지 않으므로 시스템의 성능을 향상시키는 장점을 가진다. 하지만 Jena를 이용하면 대용량의 온톨로지의 경우 로딩이 되지 않는 경우가 있으며 로딩 속도가 너무 오래 걸리는 단점이 있고, Sound하지만 Incomplete한 결과 값을 갖는다.

3. OntoThink-K[®]의 구성

지식 기반 정보유통 플랫폼 OntoFrame-K[®] 상에서 추론을 이용하여 연구자 간 협업 서비스를 제공할 수 있도록 하는 OntoThink-K[®]에 대해 설명하고자 한다. OntoThink-K[®]은 그림 1과 같이 구성된다. 본 추론 서비스 시스템은 URI 서버를 이용하여 RDF 트리플을 생성하고 전방향 추론 방법으로 규칙에 의해 트리플을 확장하며 SPARQL을 통해 서비스에 필요한 결과 값을 검색해낸다. 특히 이 모든 과정은 DBMS를 이용하여 진행하며 별도의 추론 엔진을 사용하지 않고도 정확성이 보장되는 지식을 생성 관리할 수 있다는 특징을 가진다.

국가 과학기술 기반 정보 온톨로지 상에 정의된 각종 자원 (Resource)들을 다루는 URI 서버에서 정의된 DB-클래스 매핑 정보와 DB-트리플 변환 규칙을 통해 지식 데이터들이 트리플로 변환된다. 변환된 트리플은 온톨로지 확장 모델을 거쳐 트리플 형태로 DBMS에 저장되게 된다. 그 후, 추론 서비스 인터페이스로부터 사용자가 취한 동작이나 입력한 값에 의해 자동 생성된 사용자 질의는 SPARQL 형식으로 변환된다. Persistent Model에 SPARQL로 검색 한 후 결과 값은 후처리 과정을 거쳐서 XML 형식으로 생성된 후, 추론 서비스 인터페이스에 전달되어 시각화 된다.

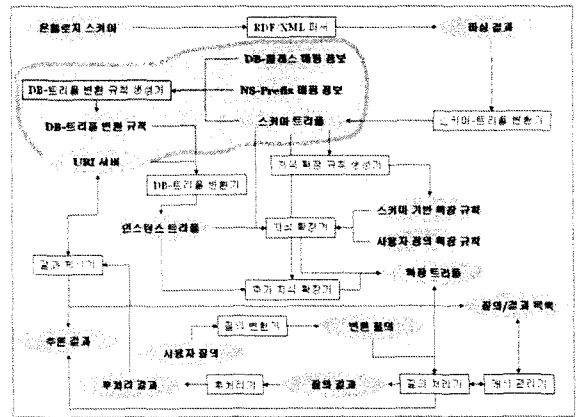


그림 1. OntoThink-K[®] 구성도

4. 온톨로지 확장 모델

4.1 온톨로지 확장 모듈의 구성

온톨로지 확장 모듈의 구성은 그림 2와 같다. OWL 온톨로지나 인스턴스 데이터는 트리플 변환기를 통해 트리플 형태 (Subject, Predicate, Object)로 변환되어 Triple DB에 저장된다. 이렇게 변환 데이터들은 단순한 1차적 트리플 형태로 추론 서비스에서 필요한 새로운 지식을 갖고 있지 않다. 따라서 본 모듈에서는 전방향 추론 방법에 의해 새로운 지식 데이터를 생성하여 확장된 지식을 이용해서 고수준의 온톨로지 추론 서비스를 제공하게 된다.

DB에 저장된 트리플들은 RDFS entailment rules를 기술한 온톨로지 스키마 규칙에 의해서 확장된다[6]. 그 후, 서비스에서 필요한 규칙들을 기술한 사용자 정의 규칙에 의해 트리플이 확장되어 Extended Triple Model에 저장된다. Persistent Model 형태로 저장된 트리플들은 추론 서비스에서 요청하는 SPARQL 질의에 의해 지식이 검색된다.

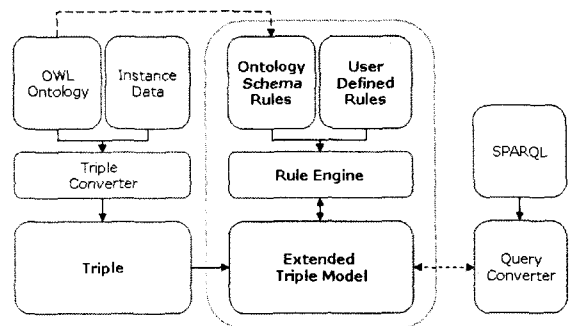


그림 2. 온톨로지 확장 모듈의 구성도

지금까지 개발된 추론 엔진들은 개발의 초기 단계로써 아직까지 **Completeness**와 **soundness**를 모두 만족하는 엔진이 없으며 OWL의 표현력을 제대로 지원해주지 않거나 대용량의 데이터를 로딩하지 못하는 등 성능 면에서 많은 문제점을 갖고 있다. 따라서 우리는 대용량의 온톨로지 데이터를 다루기 위해서 추론 엔진을 대체할만한 추론 기능을 지원할 수 있는데 데이터 저장 모델이 필요하게 되었다. 이에 따라 DBMS를 이용한 온톨로지 추론 모델을 구현하게 되었다. 이 모델은 대용량의 온톨로지를 처리할 수 있으며 지식 데이터의 **Completeness**와 **soundness**를 보장해준다.

4.2 모듈별 역할

온톨로지 확장 모델을 처리하는 모듈은 트리플을 입력받아 온톨로지 스키마 규칙과 사용자 정의 규칙을 적용하여 새로운 지식 모델로 확장하는 역할을 한다.

(1) 규칙 엔진

규칙 엔진의 역할은 입력된 트리플에 규칙을 적용하여 새로운 지식 데이터를 생성해내는 역할을 한다. 규칙 엔진은 규칙 파일을 읽어서 해석하는 부분, 규칙에 적용되는 데이터를 트리플에서 검색하는 부분, 규칙을 적용시킨 새로운 트리플을 생성하는 부분, 생성된 트리플을 확장 트리플 DB에 저장되는 부분으로 나눌 수 있다. SQL을 이용하여 데이터를 질의한 후, 결과 값에 따라 새로운 트리플을 생성하고 확장된 트리플 모델에 생성된 지식을 추가하게 된다. 그리고 본 규칙 엔진의 특징은 DBMS 기반으로 데이터를 검색, 생성, 저장한다는 것과 **negation**의 처리를 지원한다는 것이다.

(2) 온톨로지 스키마 규칙

온톨로지 스키마 규칙은 RDF Semantics에서 정의한 RDFS entailment rules에 따라 자동으로 생성된 규칙이다. 온톨로지 스키마에서 RDFS entailment가 적용되는 부분이 있으면 자동으로 규칙이 생성된다[6]. 현재 OntoThink-K®의 서비스에서 필요한 스키마 관계는 두 클래스들간의 서브 클래스 관계만 사용되고 있어서 이 규칙 파일은 그림 3의 rdfs9의 규칙이 적용되어 생성되었다.

그림 3에서 rdfs9와 같이 'rdfs:subClassOf' 라는 Predicate를 가지는 Triple이 있다면 Subject의 클래스 타입을 갖는 인스턴스 'vvv'는 Object의 클래스 타입도 갖게 된다는 것이다. 쉽게 말해서 서브 클래스 관계로 생성된 클래스들은 자식 클래스의 인스턴스들은 모두 부모 클래스의 인스턴스이므로 부모 클래스의 인스턴스로 확장될 수 있는 규칙이 생성되어야 한다.

| Rule Name | if E contains: | then add: |
|-----------|---|--|
| rdfs1 | URI aaa B. | URI ccc type ccc:Literal. |
| rdfs2 | where B is a plain literal (with or without a language tag) | where _rdfs identifies a blank node different to B by rule rdfs1 |
| rdfs3 | URI aaa xxx. | URI ccc type xxx. |
| rdfs4a | URI aaa vvv. | VV type xxx. |
| rdfs4b | URI aaa vvv. | VV type xxx. |
| rdfs5 | URI ccc:subPropertyOf VV. | URI ccc:subPropertyOf VV. |
| rdfs6 | URI ccc type ccc:Property. | URI ccc:subPropertyOf VV. |
| rdfs7 | URI ccc:subPropertyOf VV. | URI ccc:subPropertyOf VV. |
| rdfs8 | URI ccc type ccc:Class. | URI ccc:subClassOf ccc:Resource. |
| rdfs9 | URI ccc:subClassOf VV. | VV type xxx. |
| rdfs10 | URI ccc type ccc:Class. | URI ccc:subClassOf VV. |
| rdfs11 | URI ccc:subClassOf VV. | URI ccc:subClassOf VV. |
| rdfs12 | URI ccc type ccc:ConstraintProperty. | URI ccc:subPropertyOf ccc:Resource. |
| rdfs13 | URI ccc type ccc:Resource. | URI ccc:subClassOf ccc:Resource. |

그림 3. RDFS Entailment rules

국가 과학기술 기반 정보 온톨로지에서는 그림 4와 같은 클래스들의 계층 구조를 갖고 있다. rdfs9를 적용하되 중복된 데이터가 발생하지 않게 규칙 파일을 만드는 알고리즘을 이용하여 SubClass관계에 있는 클래스들을 확장하는 온톨로지 스키마 규칙을 자동 생성하였다.

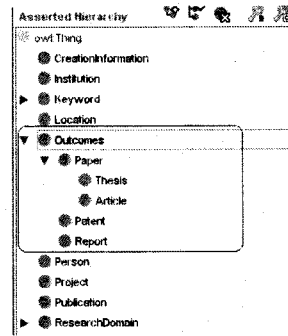


그림 4. 클래스 계층구조

(3) 사용자 정의 규칙

사용자 정의 규칙에서는 추론 서비스를 제공하기 위해 필요한 지식들을 규칙을 이용한 전방향 추론을 이용하여 미리 생성해놓는다. 단순한 트리플만으로는 찾아 낼 수 없는 정보를 이용하기 위해 사용자가 규칙을 정의하여 새로운 지식들을 생성해낸다. 본 시스템에서는 22개의 사용자 정의 규칙을 이용하여 트리플에 전방향 추론 방법을 적용하였다. 규칙의 일부를 살펴보면 그림 5와 같다.

그림 5를 살펴보면, 각 성과물마다 연도가 할당된 프로퍼티가 다르기 때문에 추론 서비스에서 성과물들의 연도 정보를 가져오기 위해서는 성과물마다 서로 다른 접근 방법을 이용하여 Predicate를 검색해야 한다. 이런 단점을 보완하기 위해서 모든 성과물의 연도는 'hasPublicationYear'라는 Predicate를 사용하도록 새로운 지식을 생성했다.

■ Article

■ x hasPublication y, y yearOfPublication z → x hasPublicationYear z

■ Thesis

■ x yearOfPublishingThesis y → x hasPublicationYear z

■ Patent

■ x yearOfApplication y → x hasPublicationYear y

■ Report

■ x hasOriginatedProject y, y endingYearOfProject z → x hasPublicationYear z

그림 5. 성과물 연도확장 규칙

이외에 성과물의 저작 정보, 성과물의 주제 분야 활동정보, 인력명, 주제/분야명, 성과물 제목 등을 서비스에서 필요로 하는 전방향 추론을 거쳐 지식으로 생성하기 위해서 사용자 정의 규칙을 정의하였다.

전방향 추론을 위해 규칙을 적용하다보면 규칙의 의존성에 따른 중복된 데이터가 만들어질 경우가 있다. 이를 방지하기 위해서 우리의 시스템에서는 규칙 적용 순서를 체계화하고 새로운 트리플을 생성하기 전에 반드시 중복 트리플을 검사하여 중복된 데이터는 생성시키지 않는다.

5. 실험 결과

현재 OntoFrame-K[®] 상에 등록된 성과물은 논문 8,786건, 특허 64건, 보고서 734건으로 모두 9,584건이다. 등록된 성과물 데이터들은 362,854개의 트리플로 변환되었다. 그리고 기관 정보, 인력 정보, 학술지 정보와 같은 성과물에 포함되지 않는 비성과물 트리플의 개수가 268,275개로 우리 시스템에서 이용되는 지식 데이터의 전체 트리플 개수는 631,158개이다.

전체 트리플들을 본 논문에서 설명한 모듈을 이용하여 온톨로지 모델로 확장해보니 비성과물 트리플은 285,982개, 성과물 트리플의 개수는 826,118개로 모두 1,112,100개의 트리플로 확장되었음을 볼 수 있다.

비성과물 트리플의 확장성이 성과물 트리플의 확장성에 비해 많지 않은 이유는 비성과물에 대한 사용자 정의 규칙이 10%밖에 되지 않기 때문이다.

그림 6은 하나의 성과물이 입력됐을 경우 생성되는 트리플(위)과 이에 따라 확장된 트리플(아래)을 보여주고 있다. 그림 6의 두 번째 테이블에서 첫 번째로 표시된 부분은 온

| subject | predicate | object |
|--|--|---|
| 1 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#Article |
| 2 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasPublication | http://www.kisti.re.kr/isri#SOU_SOP000116 |
| 3 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#koreanNameOfArticle | SALS를 이용한 고분자용액의 상전환 기구에 관한 ... |
| 4 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#CreationInformation |
| 5 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasCreationInformation | http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... |
| 6 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#hasCreator | http://www.kisti.re.kr/isri#PER_7010342497 |
| 7 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#hasInstitutionOfCreator | http://www.kisti.re.kr/isri#INS_9R9048 |
| 8 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#orderOfCreator | 1 |
| 9 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#contributionWeightOfC... | 0.5 |
| 10 http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#CategoryDomain |
| 11 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasCategoryDomain | http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... |
| 12 http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... | http://www.kisti.re.kr/isri#hasCategoryKeyword | http://www.kisti.re.kr/isri#CAT_000000 |
| 13 http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... | http://www.kisti.re.kr/isri#weightOfCategoryKey... | 0.0 |
| 14 http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#TopicDomain |
| 15 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasTopicDomain | http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... |
| 16 http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... | http://www.kisti.re.kr/isri#hasTopicKeyword | http://www.kisti.re.kr/isri#TOP_JT0005031 |
| 17 http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... | http://www.kisti.re.kr/isri#weightOfTopicKeyword | 1.0 |

| subject | predicate | object |
|--|--|---|
| 1 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#Article |
| 2 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasPublication | http://www.kisti.re.kr/isri#SOU_SOP000116 |
| 3 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#koreanNameOfArticle | SALS를 이용한 고분자용액의 상전환 기구에 관한 ... |
| 4 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#CreationInformation |
| 5 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasCreationInformation | http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... |
| 6 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#hasCreator | http://www.kisti.re.kr/isri#PER_7010342497 |
| 7 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#hasInstitutionOfCreator | http://www.kisti.re.kr/isri#INS_9R9048 |
| 8 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#orderOfCreator | 1 |
| 9 http://www.kisti.re.kr/isri#KIST11.PCD.100001_70... | http://www.kisti.re.kr/isri#contributionWeightOfC... | 0.5 |
| 10 http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#CategoryDomain |
| 11 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasCategoryDomain | http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... |
| 12 http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... | http://www.kisti.re.kr/isri#hasCategoryKeyword | http://www.kisti.re.kr/isri#CAT_000000 |
| 13 http://www.kisti.re.kr/isri#KIST11.PCD.100001_00... | http://www.kisti.re.kr/isri#weightOfCategoryKey... | 0.0 |
| 14 http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#TopicDomain |
| 15 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasTopicDomain | http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... |
| 16 http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... | http://www.kisti.re.kr/isri#hasTopicKeyword | http://www.kisti.re.kr/isri#TOP_JT0005031 |
| 17 http://www.kisti.re.kr/isri#KIST11.PCD.100001_JT... | http://www.kisti.re.kr/isri#weightOfTopicKeyword | 1.0 |
| 18 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#Paper |
| 19 http://www.kisti.re.kr/isri#KIST11.PCD.100001 | http://www.w3.org/1999/02/22-rdf-syntax-ns#type | http://www.kisti.re.kr/isri#Outcomes |
| 20 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#hasPublicationYear | 2002 |
| 21 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#createdByPerson | http://www.kisti.re.kr/isri#PER_7010342497 |
| 22 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#createdByInstitution | http://www.kisti.re.kr/isri#INS_9R9048 |
| 23 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#createdByLocation | http://www.kisti.re.kr/isri#LOC_305806 |
| 24 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#isClassifiedBy | http://www.kisti.re.kr/isri#TOP_JT0005031 |
| 25 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#isCategorizedBy | http://www.kisti.re.kr/isri#CAT_000000 |
| 26 http://www.kisti.re.kr/isri#PER_7010342497 | http://www.kisti.re.kr/isri#creatorOf | http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 |
| 27 http://www.kisti.re.kr/isri#INS_9R9048 | http://www.kisti.re.kr/isri#institutionOf | http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 |
| 28 http://www.kisti.re.kr/isri#OBJ_KIST11.PCD.100001 | http://www.kisti.re.kr/isri#nameOfOutcomes | SALS를 이용한 고분자용액의 상전환 기구에 관한 ... |

그림 6. Triple Model과 Extended Triple Model

톨로지 스키마 규칙에 의해 확장되어 생성된 지식이고 아랫쪽에 표시된 부분은 사용자 정의 규칙에 의해 확장되어 생성된 지식을 보여주고 있다.

6. 결론

본 논문에서는 OntoFrame-K[®]의 추론 서비스에 이용되는 DBMS 기반의 온톨로지 확장 모델에 대해 설명하였다. 기존의 추론 엔진들을 쓰면 대용량의 데이터를 로딩 및 추론할 수 없고 데이터 로딩 속도와 추론 성능이 좋지 않다. 그리고 추론 엔진들의 특징에 따라 지원할 수 있는 표현력도 차이가 나며 Sound하고 complete한 추론 엔진을 찾기 힘든 실정으로 우리의 시스템에 적합한 추론 엔진을 선택하기 힘들었다. 그래서 우리는 기존의 추론 엔진을 사용하지 않고 우리 시스템에서 필요로 하는 기능을 지원하는 최적화된 온톨로지 확장 모델을 설계하고 구현하였다.

OntoFrame-K[®] 시스템의 특성상 대용량의 데이터를 지원하기 위해 Persistent Model로 DBMS를 이용하게 되었으며, 온톨로지의 트리플 모델과 규칙을 이용한 전방향 추론 방법을 이용하여 지식 데이터를 확장하였다. 이 모델은 Persistent Model 형태이므로 추론을 실행할 때마다 온톨로지를 로딩할 필요 없이 처음 데이터를 저장할 때 데이터를 추론하여 확장하기 때문에 서비스 이용시 데이터의 접근 성능을 높여준다.

전방향 추론 방법의 단점으로 데이터양이 많아질수록 시스템에 많은 부하를 가져오게 되는데, 우리 시스템은 인스턴스 데이터가 등록될 때마다 해당되는 트리플에 대해서만

규칙을 확장하기 때문에 데이터 저장시 성능 면에서는 큰 문제를 갖지 않는다. 다만 DBMS를 이용하기 때문에 데이터가 많아질수록 SPARQL을 이용한 검색 시 속도가 느려지는 문제가 발생할 수 있다. 그래서 우리는 이 문제를 해결하기 위해 Predicate에 따라 테이블을 분리하여 트리플을 관리하고, 캐시를 이용하는 방법을 통해 서비스 수행 속도 향상에 포커스를 맞춰 시스템을 보완하고 있다.

참고 문헌

- [1] 성원경, 정한민, 박동인, "OntoFrame-K[®]: 협업 연구 지원을 위한 시멘틱 웹 기반 지식정보 공유, 유통 플랫폼", 정보과학회지, 제 24권 제4호, 2006. 4.
- [2] 이미경, 정한민, 이승우, 성원경, "국가 과학기술 R&D 기반정보 온톨로지 구축 및 적용", 제25회 한국정보처리학회 춘계학술발표대회, 제13권 제1호, 2006. 5
- [3] J. Broekstar, A. Kampman, F. Harmelen, "Sesame : A Generic Architecture for Storing and Querying RDF", Published at the International Semantic Web Conference 2002, Sardinia, Italy.
- [4] 김기성, 유상원, 이태휘, 김형주, "RDF 스키마 함의 규칙 적용 순서를 이용한 RDFS 추론 엔진의 최적화", 정보과학회지논문지 : 데이터베이스 제33권 제2호, 2006. 4
- [5] "Jena2 ontology API", <http://jena.sourceforge.net/ontology/index.html>
- [6] "RDF Semantics", W3C Recommendation 10 February 2004, <http://www.w3.org/TR/rdf-ml/>