Y-HisOnto: Q&A 시스템에서의 활용을 위한 역사 온톨로지 모형

이인근O, 정재은, 황도삼
경북대학교 의료정보학과O, 영남대학교 컴퓨터공학과
inkeunlee@gmail.comO, j2jung@gmail.com, dshwang@yu.ac.kr

Y-HisOnto: A History Ontology Model for Q&A System

In Keun LeeO, Jason J. Jung, Dosam Hwang
Department of Medical Informatics, Kyungpook National University School of MedicineO
Department of Computer Engineering, Yeungnam University

요 약

본 논문에서는 시간 개념이 포함된 역사적 지식을 표현할 수 있는 사건 온톨로지(event ontology) 기반의 역사 온톨로지 모형인 Y-HisOnto 를 제안한다. 제안한 역사 온톨로지 모형은 기존의 온톨로지에서 사용되는 이진 관계(binary-relationship)로 표현된 단편적 지식들을 조합하여 다진 관계(n-ary relationship)를 이용하여 역사적 사건 관련 지식을 표현한다. 제안한 온톨로지 모형에 기반하여 사건 중심의 지식을 온톨로지로 구축하고, 사건 관련 질의에 대해 온톨로지 논리 검색 실험을 수행함으로써 제안한 온톨로지 모형이 Q&A 시스템에서 효과적으로 활용될 수 있음을 확인한다.

주제어: 사건 온톨로지(event ontology), 역사 온톨로지, 질의응답 시스템(Q&A System)

1. 서론

컴퓨터를 통한 지식의 구축 및 활용 요구에 따라 질의응답(Question and Answer: Q&A) 시스템에 관한 연구가 수행되어 왔고[1-3], 컴퓨터를 통한 지식의 표현을 위해 최근에는 온톨로지를 이용하는 방법에 대한 연구가 활발히 수행되고 있다[4-6]. T.Gruber의 온톨로지 정의에 의하면, 온톨로지는 광범위한 지식 보다는 관심 영역 내의 합의된 지식을 체계적으로 구축하는 것이 온톨로지의 활용 측면에서 더 의미가 있다고 해석할 수 있다. 이에 따라 다양한 분야에서 분야 전문 지식을 온톨로지로 구축하고 효과적으로 활용하기 위한 연구가 수행되어 왔다. 특히, 유럽에서는 VICODI project[7]를 통해 유럽의 역사적 지식을 온톨로지로 표현하려는 연구가 수행되었다. 타 분야의 지식과 달리, 역사적 사실의 전개는 주로 사건(event)에 기반하여 기술되고 있으며, 사건은 시간에 종속되어 표현된다. 따라서 이러한 복잡한 형태의 역사적 지식을 표현하고 활용하는 방법의 연구가 필요하다.

역사에서 하나의 사건은 주체, 시간, 대상, 장소, 행위 등과 같이 개념들이 복잡한 의미 관계를 형성하고 있으며, 사건이 일어나는 원인이나 특정 사건에 의해 발생되는 결과의 표현을 위해서는 역사적 지식이 사건 중심으로 표현되어야 한다. 실제 RDF(S), F-logic, OWL (web ontology language)등과 같은 다양한 온톨로지 언어[4]들이 지식의 표현을 위해 사용되고 있고, 이들 언어들은 이진 관계(binary-relationship)로 단편적 지식을 표현한다. 특히 OWL[8]에서는 지식의 표현을 위해 개념(class), 개체(individual), 속성(property), S-P-O (subject-predicate-object) 형태의 개념간의 관계(relationship), 개념의 범위를 지정하기 위한 제약사항(restriction), 그리고 string, integer 등과 같은 데이터(data) 등의 온톨로지 요소를 정의하고 있다. 그러나 OWL에서와 같이 단순한 이진관계로는 복잡한 사건 지식을 표현하기 어렵다.

따라서 본 논문에서는 시간 개념이 포함된 역사적 지식을 사건 기반으로 구축하고 활용할 수 있는 온톨로지 모형인 Y-HisOnto 를 제안한다. 제안한 온톨로지 모형은 이진 관계를 조합하여 다진 관계(n-ary relationship)를 표현하는 사건 온톨로지(event ontology)[5,6]를 이용한다. 그리고 제안한 온톨로지 모형에 기반하여 사건 중심의 지식을 구축하고, 제안한 온톨로지의 Q&A 시스템에서의 효용성을 확인하기 위해 온톨로지 논리 검색 실험을 수행한다.

2. 역사 온톨로지 모형의 설계

이진관계를 이용하여 이벤트 중심의 역사 지식을 표현하기 위해서 그림 1과 같은 형태로 온톨로지를 구성한다. 즉, 특정 "사건" 개념을 중심으로 사건의 부가적인 정보를 표현하기 위해, 사건을 유발한 "인물", 사건이 발생한 "장소", 사건의 발생 및 종료 시간, 그리고 사건의 발생 목적과 해당 사건으로 인해 야기되는 "사건"등과 같은 정보가 필요하다. 그림 1에서 보는 바와 같이, 이러한 정보를 온톨로지로 표현하기 위해 "인물", "장소", "사건"을 개념으로 표현하고, 화살표 근처의 "주체", "장소", "인과", "목적"은 모두 개념 간의 관계 형성을 위한 속성으로 나타낸다. 특히, "시작시간"과 "종료시간"의 속성은 사건의 발생 및 종료 시점을 나타내는 것으로, 그 값은 정수(integer)로 표현하도록 하였다. 그러나 이 값은 날짜

(datetime) 등과 같은 다양한 자료형으로 정의할 수 있다.

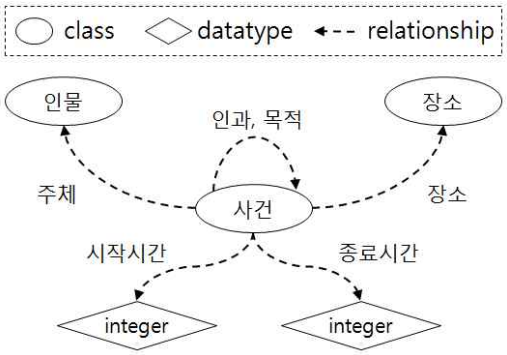

그림 1. 사건 지식 표현을 위한 온톨로지 스키마 예

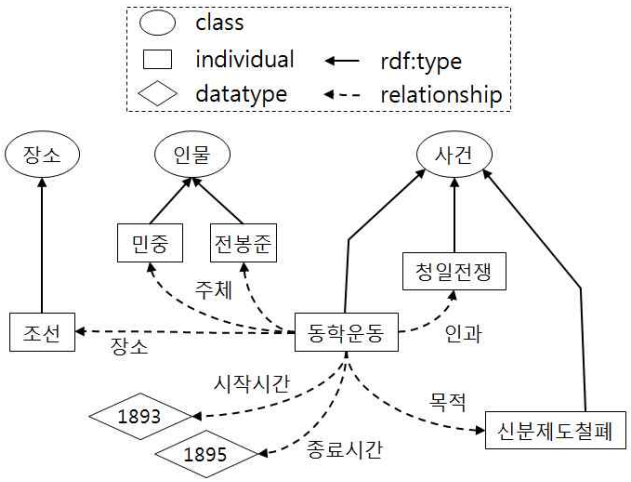

그림 2. 역사적 사건 지식의 온톨로지 표현 예

그림 2는 "전봉준을 중심으로 신분제도철폐를 위해 조선에서 일어난 민중 봉기인 동학운동은 청일전쟁의 원인이 되었다"의 역사적 사건 지식을 그림 1의 온톨로지 스키마에 기반하여 온톨로지로 표현한 예를 보인다. 즉, [표현 1]에서와 같이 이진관계들을 조합으로 하나의 사건 지식은 상세하게 나타낼 수 있다. [표현 1]에서는 일차 논리(First-order logic)를 이용하여 P(S, O) 형태로 사건 지식을 나타낸 것이다.

[표현 1]

주체(동학운동, 민중)	주체(동학운동, 전봉준)
장소(동학운동, 조선)	인과(동학운동, 청일전쟁)
시작시간(동학운동, 1893)	종료시간(동학운동, 1895)
목적(동학운동, 신분제도철폐)	

3. 역사 온톨로지 모형의 활용

제안한 온톨로지 모형에 기반하여 특정 역사적 사건 지식을 Protégé[9]를 이용하여 온톨로지로 구축하고, 구축한 온톨로지를 Q&A 시스템에서의 활용에 대한 효용성을 확인하기 위해 SPARQL[10]을 이용하여 몇 가지 질의에 대한 온톨로지 논리 검색을 수행한다. 그림 3은 온톨로지에 정의한 개체(individual)를 보이며, 그림 4는 개체간의 관계를 통해 사건 지식을 표현한 예를 보인다. 또한, 그림 5는 Protégé를 이용하여 역사 온톨로지를 구축하는 것을 보인다. 본 논문에서 온톨로지 요소에 대한 Namespace는 "http://ontology.yu.ac.kr/event"이며, 이에 대한 Prefix를 "ynu"로 지정하였다. 그러나 그림 3과 4에서는 편의상 이를 생략하였다.

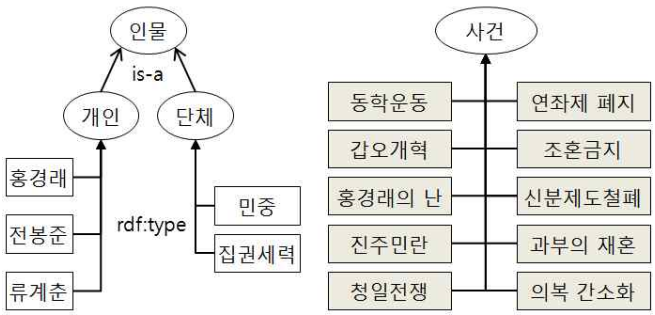

그림 3. 역사 온톨로지에서의 개체 정의

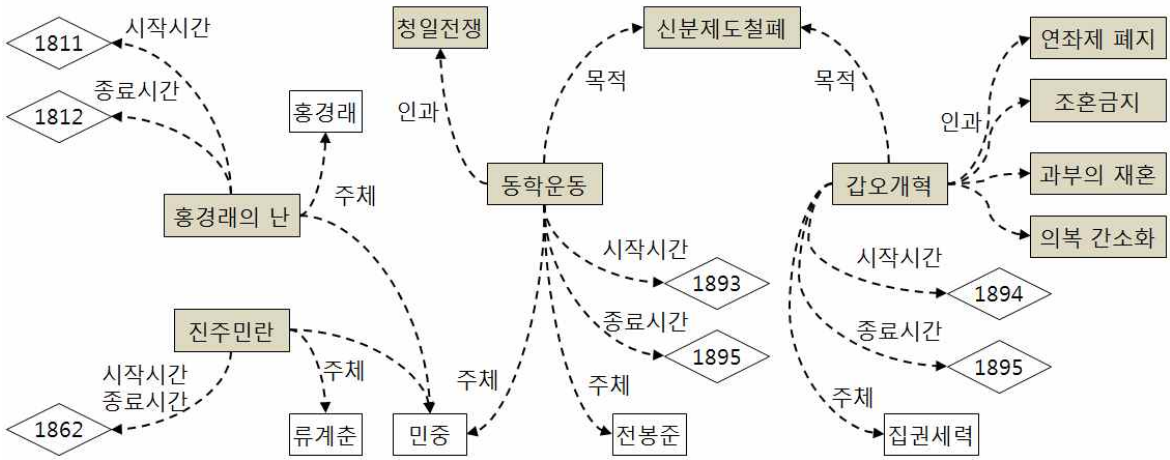

그림 4. 역사 온톨로지 구축 예

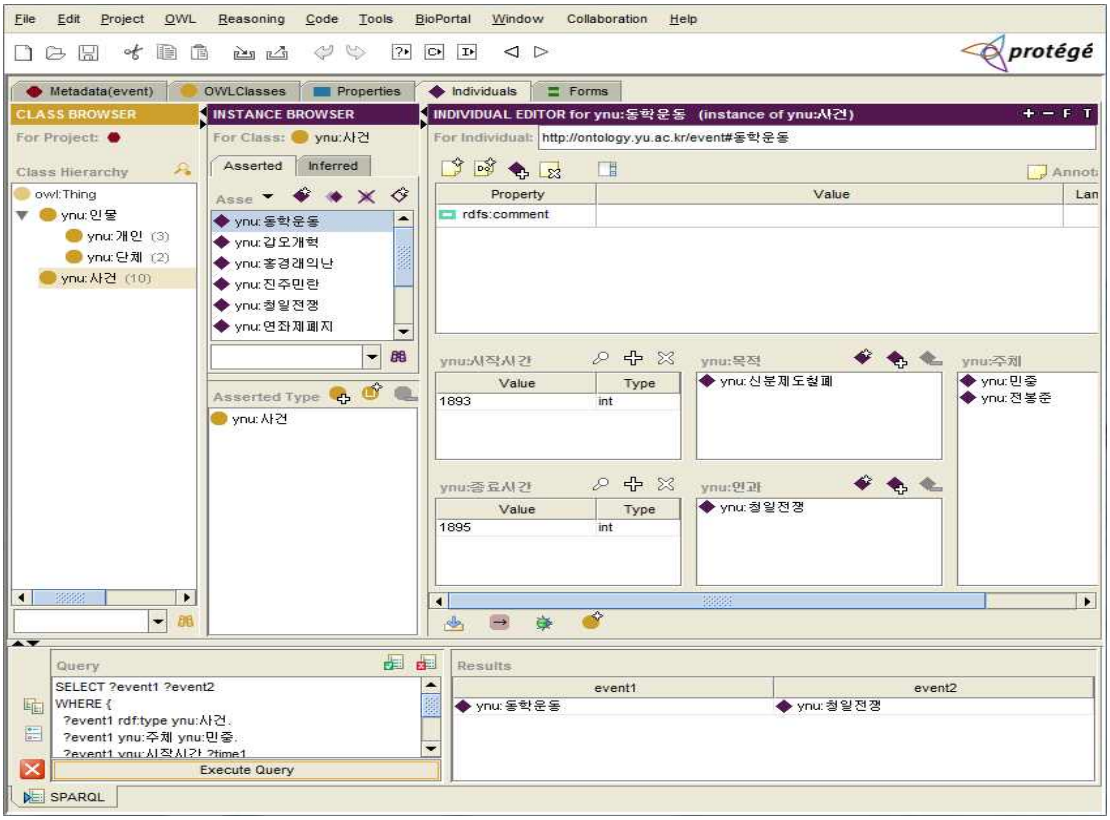

그림 5. Protégé를 이용한 역사 온톨로지 구축

특정 Q&A 시스템에서 구축한 역사 온톨로지를 이용하여 사용자의 3 가지 질의에 대해 답을 제시하는 상황을 가정하여 실험을 수행하였다. 온톨로지의 논리 검색은 Protégé에서 제공하는 검색엔진 및 검색 인터페이스를 이용하였다. 논리 검색을 위한 사용자의 3 가지 질의는 다음과 같다.

[질의1] "민중"이 일으킨 사건은 무엇인가?
[질의2] "갑오개혁"으로 인해 야기된 사건은 무엇인가?
[질의3] 고종의 재위기간(1863~1897) 중 일어난 "민중" 사건(1)과 그로 인해 야기된 사건(2)은 무엇인가?

[질의1]로부터 주체가 "민중"인 사건의 검색을 위해 [SPARQL 질의 1]을 생성하여 논리검색을 수행하였고, 그림 6과 같이 검색 결과로 "홍경래의난", "동학운동", "진주민란"을 얻었다.

[SPARQL 질의 1]

```
SELECT ?event
WHERE {
  ?event rdf:type ynu:사건.
  ?event ynu:주체 ynu:민중.
}
```

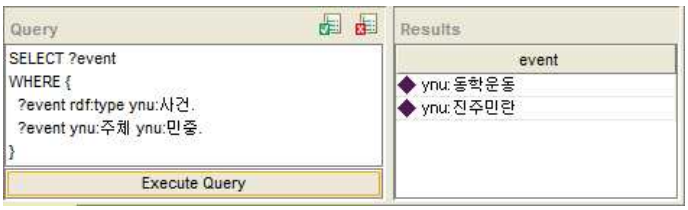

그림 6. [질의1]에 대한 온톨로지 논리 검색

[질의2]로부터 사건 "갑오개혁"이 원인이 되는 사건의 검색을 위해 [SPARQL 질의 2]을 생성하여 논리검색을 수행하였고, 그림 7과 같이 검색 결과로 "과부의 재혼", "조혼금지", "연좌재폐지", "의복간소화"를 얻었다.

[SPARQL 질의 2]

```
SELECT ?event
WHERE {
  ynu:갑오개혁 ynu:인과 ?event.
}
```

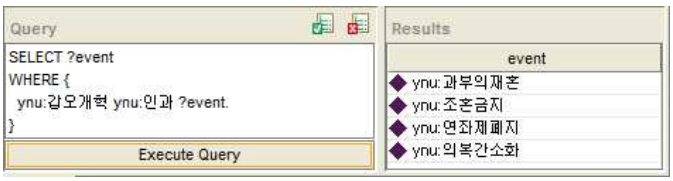

그림 7. [질의2]에 대한 온톨로지 논리 검색

[질의3]으로부터 1863년~1897년 사이에 발생한 사건으로써 주체가 "민중"인 사건(1)의 검색, 그리고 사건(1)과 인과관계가 있는 사건(2)의 검색을 위해 [SPARQL 질의 3]을 생성하여 논리검색을 수행하였다. 검색 결과를 통해 그림 8과 같이 사건(1)은 "동학운동"을, 사건(2)는 "청일전쟁"임을 알 수 있었다.

[SPARQL 질의 3]

```
SELECT ?event1 ?event2
WHERE {
  ?event1 rdf:type ynu:사건.
  ?event1 ynu:주체 ynu:민중.
  ?event1 ynu:시작시간 ?time1.
  ?event1 ynu:종료시간 ?time2.
  FILTER ( ?time1 >= 1863 && ?time2 <= 1897  )
  ?event1 ynu:인과 ?event2.
}
```

그림 8. [질의3]에 대한 온톨로지 논리 검색

본 실험을 통해 역사적 사건 지식을 온톨로지로 구축하고, Q&A 시스템에서 복잡한 사용자의 질의에 대해 논리적 검색을 통해 답을 제공할 수 있음을 확인하였다.

4. 결론

본 논문에서는 복잡하게 표현되는 역사적 사건을 온톨로지 모형인 Y-HisOnto를 설계하여 제시하고 구축함으로써 Q&A 시스템에서의 활용 가능성을 확인하였다.

구축한 온톨로지의 활용을 위해서는 온톨로지의 검색을 통해 지식을 추출해야 한다. 또한 검색 대상과 온톨로지 모형의 설계에 따라 검색 방법이 달라질 것이다. 즉, 본 실험에서 사용한 SPARQL 질의는 실험을 위해 구축한 역사 온톨로지를 기반으로 작성한 것이다. 따라서 만일 온톨로지 스키마가 새롭게 정의되고, 그에 기반하여 새롭게 온톨로지가 구축된다면 좀 더 다양한 조건의 검색질의가 생성되어야 한다.

역사적 지식을 정형화하여 표현함에 있어 다음과 같은 문제들을 고려해야 한다. 즉, (1)동일한 사건 및 대상의 상태는 시간에 따라 변화할 수 있고, (2) 역사적 기록은 주관적이어서 온톨로지에서의 지식 표현의 일관성의 유지가 어려우며, (3)역사적으로 복잡한 사건 간의 인과관계를 단편적으로 정의하기가 어렵다. 따라서 본 논문에서 제시한 온톨로지 스키마는 모든 역사적 지식을 구축하기에는 무리가 있다.

본 논문에서는 시간 개념을 포함된 사건 지식 및 그들 사이의 인관관계를 형성하기 위한 방법을 제안하고, 그의 활용 가능성을 확인하였다는 것에 의미가 있다고 본다. 또한, 역사 지식의 정형화 및 온톨로지로의 표현 방법에 대한 구체적인 연구는 지속적으로 수행되어야 하며, 이에 대한 연구는 향후 지속할 계획이다.

Acknowledgement

본 연구는 미래창조과학부 및 한국산업기술평가관리원의 산업융합원천기술개발사업의 일환으로 수행하였음 [10044457, 자율지능형 지식/기기 협업 프레임워크 기술개발]

참고문헌

[1] S. Kim, D. Baek, S. Kim, and H. Rim "Question Answering Considering Semantic Categories and Co-occurrence Density," In Proceeding of 9th Text Retrieval Conference, pp. 317-325, 2000.
[2] K.C. Litkowski, "CL Research Experiments in TREC-10 Question Answering," In Proceeding of 10th Text Retrieval Conference, pp. 122-131, 2001.
[3] S. Vassiliadis, G. Triantafyllos, and W. Kobrosly, "A Fuzzy Reasoning Database Question Answering System," IEEE Transactions on knowledge and data engineering, Vol. 6, No. 6, pp. 868-882, 1994.
[4] A.G. Perez, M.F. Lopez, and O. Corcho, Ontological Engineering, Springer, 2005.
[5] H. Lin and J. Liang, "Event-based ontology design for retrieving digital archives on human religious self-help consulting," In Proceeding of IEEE International Conference on EEE'05, pp. 522-527, 2005.
[6] Y. Raimond and S. Abdallah, "The Event Ontology," http://motools.sourceforge.net/ event/event.html, 2004.
[7] VICODI project, http://www.vicodi.org, 2013.
[8] OWL Web Ontology Language Guide, http://www.w3.org/TR/owl-guide, 2009.
[9] Protégé, http://protege.stanford.edu, 2013.
[10] SPARQL Query Language for RDF, http://www.w3.org/TR/rdf-sparql-query, 2013.