

Bossam 추론 엔진을 이용한 한의학 온톨로지 개발

문경실*, 박수현**

*동서대학교 일반대학원

**동서대학교 컴퓨터정보공학부

e-mail:siri1210@hanmail.net

Development of Oriental Medical Ontology using Bossam Inference Engine

Kyung-Sil Moon*, Su-Hyun Park*

*Graduate School, General Dongseo University

**Department of Computer Engineering, Dongseo University

요 약

의료 분야의 정보화 움직임으로 인해 단순한 정보 저장과 검색 시스템에서 벗어나 지능화된 서비스를 제공 해주는 시멘틱 웹 기반의 의료 시스템이 요구되고 있다. 이에 한의학 분야도 한의사의 진단을 보조할 수 있는 지식 기반 시스템에 대한 요구가 증대되고 있으며 관련 시스템들이 개발되어지고 있다. 온톨로지는 시멘틱 웹의 핵심적인 지식 체계로 지식의 처리와 추론이 가능하므로 질의 및 논리 추론을 통하여 진단을 내리는 한의학 지식 데이터베이스 구축에 적합하다.

본 연구에서는 한의학 분야의 지식을 보다 의미적이고 체계적으로 표현하고 온톨로지를 이용한 검색 결과의 이점을 보여주기 위해 추론 기술을 접목시켜 한의학 온톨로지를 개발하였다.

1. 서론

의학 분야에서 진단과 처방, 그리고 치료 과정은 대부분 전문가의 의사결정을 통하여 이루어진다. 이러한 의학 분야에서도 IT도입을 통한 급속한 정보화가 진행되면서 의사결정과정에도 도움을 주는 진단보조시스템, 처방관리시스템과 같은 웹 기반의 지능형 시스템이 개발되어지고 있다.

이들은 단순한 정보저장과 검색 기능을 넘어서 컴퓨터 스스로가 정보의 의미를 이해하고 조작하여 정보와 지식을 처리하고 새로운 지식을 추론해 내는 시멘틱 웹을 기반으로 한다. 지식 기반의 지능형 시스템은 그 기반이 되는 지식 데이터베이스를 필요로 하는데 그것이 시멘틱 웹의 핵심적인 지식체계이자 요소인 온톨로지이다.

특히 한의학은 의학의 특성상 규격화하기 어려운 진단 조건과 객관화하기 어려운 진단 결과로 한의사의 주관적인 판단과 임상적 지식을 통한 다양한 진단 결과들이 비교적 강하게 제시된다. 하지만 많은 양의 임상지식이 쌓이고 있음에도 불구하고 체계적인 정리와 전산화가 이루어지지 않아 이에 대한 객관화와 정리된 지식 체계의 요구가 증대되고 있는 실정이다. 이러한 문제를 해결하기 위해 의료 지식을 체계화하려는 노력과 함께 진단을 보조해줄 수 있는 시스템, 그리고 시스템의 기반이 되는 지식 베이스인 온톨로지의 구축이 함께 진행되고 있다.

흔히 한의학 분야에서의 온톨로지는 질병, 증상 등 정보를 체계적이고 논리적으로 표현하고 관계를 설정함으로써 의료정보의 효과적 제공 및 의료지원이 가능한 지식 자원

이다. 또한, 한의학 온톨로지는 체계화 된 내용과 추론 기술이 더해지면서 기존의 데이터베이스보다 발전되어진 형태의 지식베이스가 된다.

따라서 본 연구에서는 기 구축된 한방 온톨로지를 확장하고 추론 기술을 접목시켜 부분적이지만 한의사의 임상 경험을 바탕으로 Bossam 추론엔진을 이용한 한의학 온톨로지를 개발한다.

2. 관련연구

2.1 Bossam 추론 엔진

온톨로지 개발을 위하여 본 논문에서는 국내에서 개발한 Bossam 추론 엔진을 사용하였다. Bossam은 규칙과 온톨로지를 모두 처리할 수 있게 만든 규칙 기반의 온톨로지 추론엔진으로 ETRI에서 2003년 초기 버전을 개발한 것을 시작으로 2005년 OWL(Web Ontology Language) 추론 기능을 확장하고 Rule-ML(Rule Markup Language) 및 SWRL(Semantic Web Rule Language) 지원이 가능하고 Java 객체 연동 기능을 확장한 v0.7,v0.8을 개발하였다. 또한, RDF-S, OWL, Rule-ML, SWRL 등의 웹 온톨로지 언어 및 웹 규칙 언어를 지원하며 온톨로지와 규칙을 결합하여 활용 가능하며 OWL의 모든 어휘를 지원한다[1].

2.2 의학 온톨로지

국내에서 한의학 분야의 온톨로지는 여러 곳에서 연구·개발 이루어지고 있으나 대부분이 소규모로 실험적인 수

준으로 이루어지고 있다.

박경모(2003)는 Protege를 이용한 한의학의 구조화된 증상 입력을 위한 온톨로지 개발에서 의학 영역에서의 계산 가능한 지식체계 즉, 온톨로지의 필요성을 강조하고 기존에 구축된 의학 온톨로지의 문제점을 제시하였다. 그리고 한의학 분야의 온톨로지를 개발하기 위한 방법론을 제안하고 간단한 한의학 증상 온톨로지 구축하였다[2].

이현실(2003)은 온톨로지 기반 한의학 처방 지식관리시스템 설계에 관한 연구에서 한의학 처방 지식관리시스템 설계에 요구되는 한의학 처방지식 온톨로지를 개발하기 위한 방법을 제시하고, 지식 추론이 가능한 한의학 처방 지식관리시스템을 제안하였다[3].

한의학 온톨로지에 관한 국외 연구는 거의 발표된 바가 없으며 현재 계속해서 진행 중인 연구 분야이다. 따라서 대표적인 의학 온톨로지에 대해서 살펴본다.

대표적인 의학 온톨로지로는 GALEN, ABFP, Digital Anatomist Ontology이 있다. GALEN은 Description logic 기반의 GRAIL (GALEN : Representation and Integration Language)에 의해서 표현된 온톨로지로서 맨체스터 대학을 중심으로 하는 유럽 연합에서 개발되었다[4]. ABFP ontology는 건강 상태, 증상들, 처치 등에 대한 추상적 기술을 제공하고 있으며 임의의 수만큼 인스턴스 (Instance)를 생성할 수 있게 되어 있다[2]. 마지막으로 Digital Anatomist Ontology는 해부학 정보에 대한 프레임 기반의 온톨로지이다[5].

3. 한의학 온톨로지 개발

3.1 한의학 온톨로지 구현 절차



(그림 1) 온톨로지 구현 절차

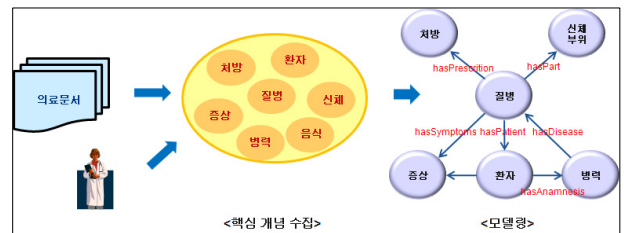
1단계 온톨로지 구현 목적 및 확인 단계에서는 온톨로지 구현 범위를 결정하며, 2단계 온톨로지 전체 구조 설계에서는 네임스페이스 선언, 3단계 온톨로지 자료 수집 및 분석에서는 한의학 용어 수집 및 용어 정의, 동의어 추출, 개념화를 완료한다. 그 후 4단계 온톨로지 내부 구조 설계에서는 추출된 용어의 계층화 작업과 관계 정의 및 제약사항을 정의한다. 그리고 5단계에서는 추론을 통한 본격적인 온톨로지를 구현하게 된다. 마지막으로 구현된 온톨로지는 지속적인 검증 및 보완 과정을 통해 완성도를 높인다.

3.2 한의학 온톨로지 설계

한의학의 도움을 받아 한의학에서의 질병분류, 질병정보, 증상표, 문진표, 건강력, 문진표, 약장, 진단정보입력과 관련된 자료를 수집하였다. 질환 개념 204개, 증상 개념 303개를 시작으로 만성질환 97개까지 총 604개의 데이터를 수집 하였다. 그 중에는 일반질병과 만성질환 모두에 포함되는 질병도 있었다. 이렇게 수집된 한의학 데이터는 분류체계를 거쳐 각각의 자료로써 전산화 작업을 통해 한의학 자원 데이터베이스로 만들어진다.

전문가로부터 받은 한의학 자료는 표준화 된 용어가 아니었고 또한 정리되지 않은 개인의 임상 지식이어서 기계가 이해하고 모든 사용자가 동일하게 이해할 수 있는 표현으로 바꾸는 부분이 문제가 되었다. 이러한 문제를 해결하기 위해 온톨로지 개발자의 한의학 분야 개념 정립 과정이 필요하였기 때문에 모든 작업은 수작업으로 개념을 추출하였다.

추출 과정은 상향식으로 수집한 자료에서 핵심용어를 선정·분류하여 개념화를 시킨 후 계층화 작업을 하였다. 계층관계는 Is-a, part-of, instance-of 등의 관계를 기반으로 구현하였다.



(그림 2) 한의학 온톨로지 모델링

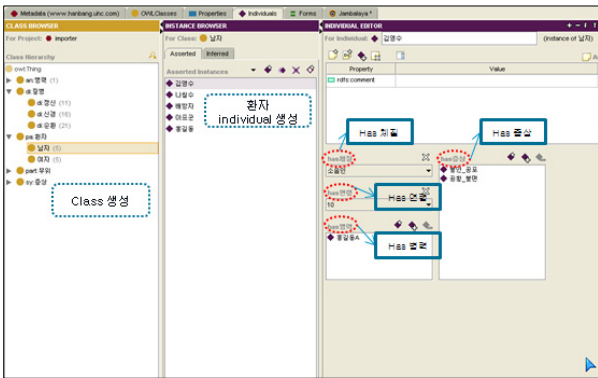
추출된 용어들은 상·하위 관계에 따라 클래스를 추가시키고 각각의 세부 클래스 및 individual을 확장하였다. 모든 클래스 및 individual의 생성이 마무리 되면 질병, 증상, 처방, 환자, 병력, 부위 등 추출 개념 간의 관계 설정 및 제약사항을 부여한 후 의료 전문가에게 질병, 증상 등 개념 및 개념 관계 설정 검증을 하였다.

3.3 한의학 온톨로지 구현

설계 단계에서 한의학 분야의 관련 용어를 추출하고 개념화하는 작업을 수행하였다. 한의학 온톨로지 구현을 위해 Protege 3.3.1 도구를 사용하여 한의학 분야의 질환(질병)을 중심으로 그에 따른 증상, 환자, 부위, 병력, 처방을 정의하고 속성들의 관계를 연결하고 제약사항을 부여하는 작업을 수행하였다. 각각의 생성된 최상위 클래스들은 Is-a 관계의 하위 클래스와 그 하위에 속하는 individual 그리고 각각의 속성을 가지고 있다.

환자는 기본적으로 남자 여자로 세분화 되며, 각각의 남자, 여자 환자는 연령(has연령), 병력(has병력), 증상(has증상), 체질(has체질) 등을 속성으로 가지고 있다. 또한 병력은 가족병력, 지병 등의 individual로 세분화 되고 가족

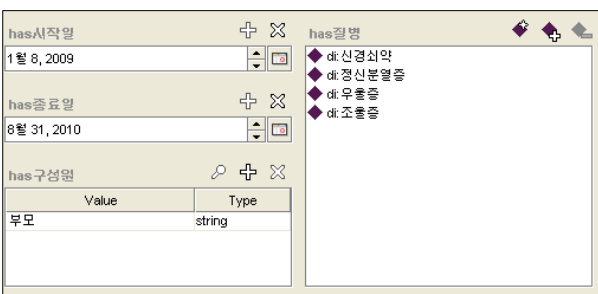
병력이 있는 환자의 경우 질병, 질병의 시작일과 종료일, 해당구성원 등의 속성을 가지고 있어 진단 시 도움이 되는 결과 값을 산출할 수 있도록 하였다.



(그림 3) protege를 이용한 온톨로지 구현

has증상 속성은 Domain으로 질병과 환자를 Range로 증상을 가지는 것과 같이 증상이 나타남에 있어 한 곳 이상의 부위를 가진다는 속성 설정, 병력은 질병을 가지고 환자는 병력을 가지는 등 다소 복잡하지만 연관되는 속성 설정으로 온톨로지의 완성도를 높이고자 하였다.

또한, 세부 질병에서 나타날 수 있는 증상들을 각각의 질병과의 관계를 정의하고 제약사항을 설정하였는데 환자는 연령 속성을 가지나 최대 한 가지 연령만을 설정할 수 있고 체질 또한 한 가지 체질만을 가질 수 있으며 병력 또한 과거에 앓았던 병의 시작일과 종료일을 설정할 수 있다. 이러한 제약사항 설정을 통하여 관계가 복잡한 형식을 정의할수록 풍부한 내용을 가진 온톨로지가 된다. 이 과정에서 한 종류의 질병이 일반 증상뿐만 아니라 또 다른 질병을 증상으로 가지고 있는 경우도 있었다.



(그림 4) 온톨로지 속성 값 부여

기본적으로 각 신체부위별 모든 질병은 그에 대한 증상, 부위, 처방 등을 가지고 있으며, 증상은 증상이 나타나는 시기, 빈도 그리고 환자는 병력을 속성으로 가지고 있어 환자의 질병 추론에 있어 정확성을 더하고자 하였다. 이는 그들의 관계를 어떻게 정의, 연결하는가에 따라 한방 온톨로지의 완성도를 높일 수 있는 것이다. 이런 속성 및 관계의 복잡성이 의하는 것은 원하는 정보를 선택하고 컴퓨터가 온톨로지라는 지식체계를 이용하여 그 정보를 찾을 때

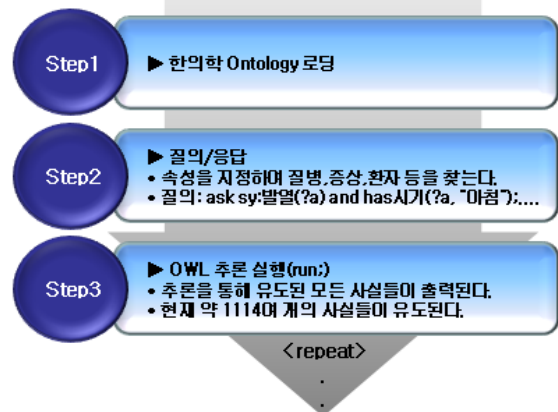
용어의 선택에 따라 다른 답변이 나올 수 있다는 것이다. 이렇게 구현된 한의학 지식은 추론 엔진을 이용한 추론 과정을 거치게 된다. 온톨로지 추론을 위하여 Bossam 추론엔진과 Bossam 추론엔진의 규칙 언어(rule language)로 Buchingae언어를 사용하였다. 아래 그림5는 로드된 OWL 파일로부터 질의문을 작성하는 단계를 보여주고 있다.

```
ask 남자(?a);
ask 여자(?a);
ask sy:발열(?a) and has시기(?a, "아침");
ask pa:환자(?x) and has증상(?x, 고온);
ask 고온 환자(?x);
ask di:호흡(?a) and has증상(?a, 감기_발열) and has시기(?a, "아침");
ask di:정신(?x) and has증상(?x, ?y);
ask di:호흡(?x) and has증상(?x, 감기_오한);
ask di:질병(?x) and has증상(?x, 공황_불면);
ask di:호흡(?x) and (has증상(?x, ?y) and has부위(?y, 부위1));
ask di:호흡(?x) and has증상(?x, 감기_발열) and has부위(?x, part:폐);
ask di:질병(?a) and has증상(?a, 복통) and has부위(?a, 우측복부) and has시기(?a, "밤") and has빈도(?a, "자주");
:
:
```

(그림 5) Bossam rule query 생성

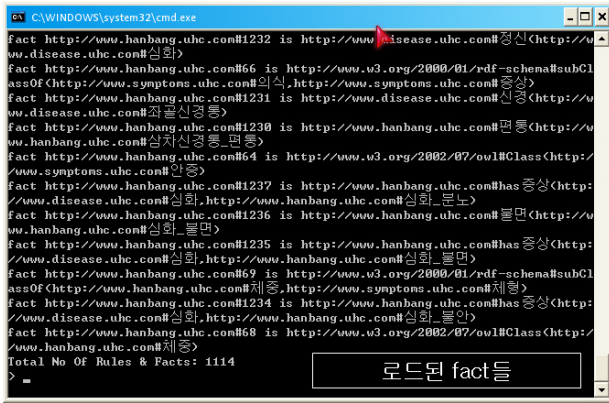
Bossam 추론 엔진은 fact들을 만들고 규칙을 정의한 후 질의문을 작성한다. 하지만, 본 논문에서는 한의학 온톨로지의 규칙을 Protege를 이용하여 정의하고 규칙을 만들었기 때문에 저장한 OWL파일을 불러와 추론하는 과정을 진행하였다.

그림6의 과정과 같이 추론은 Protege로 작성한 한의학 온톨로지 파일을 로드한 후 각각의 속성을 지정하여 그 속성에 부합되는 질병, 증상, 환자 등을 찾고 OWL 추론을 실행하면 유도된 모든 사실들이 출력된다. 다음 단계로 Step02, Step03 단계를 반복적으로 수행하면서 추론을 한다.



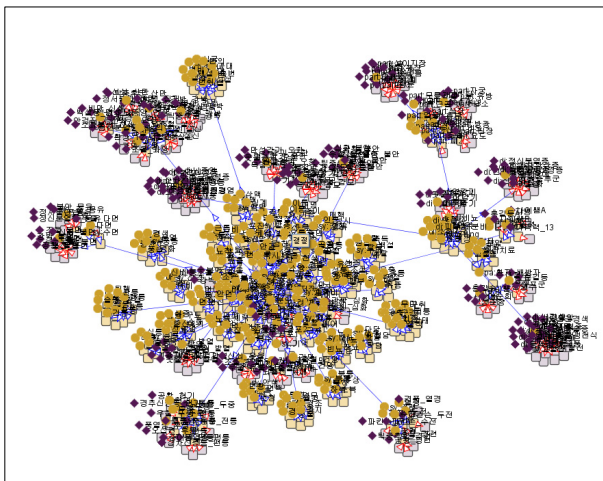
(그림 6) Bossam 추론 과정

아래 그림7은 Bossam 추론엔진을 실행하는 화면이다. 로드된 fact들을 모두 확인할 수 있으며 앞에서 정의한 규칙을 이용하여 조건 지정 후 질의문을 보내면 결론이 도출된다. 추론을 하면 약 1114여개의 사실들이 유도된다. 계속적으로 fact의 추가와 규칙의 정의를 통한 질의·응답을 실시함으로써 추론을 통한 다양하고 정확도 있는 추론 결과를 얻을 수 있다.



(그림 7) Bossam 추론 실행

현재, 모든 한의학 온톨로지 지식에 대한 모든 추론 규칙이 만들어진 상태는 아니지만 추론엔진을 이용한 추론의 기본적인 방식을 제시하고 계속해서 fact를 추가하는 작업을 하고 있기 때문에 rule이 추가로 정의 될수록 정확성 높은 결과 값과 함께 넓은 범위의 온톨로지가 될 것이다. 또한, rule 추가 작업을 반복적으로 수행하면서 한의학 진단 시 문제되는 구조적인 문제를 해결함으로써 언제든지 확장 가능한 지식베이스로서의 면모를 갖추게 된다.



(그림 8) Protege 계층도 및 관계 설정도

4. 결론

본 연구에서는 웹 온톨로지 언어(OWL)를 이용한 한의학 온톨로지를 구현하고 온톨로지를 이용한 검색 결과의 이점을 보여주기 위해 추론 기술을 접목시켜 온톨로지를 구

현하였다. 구현된 한의학 온톨로지는 한의학 분야의 정보 시스템에 활용될 수 있지만, 한의학 분야의 진단을 보조하는 것은 현실적으로 어려움이 많은 것이 사실이다. 이는 경험에 의한 임상 데이터의 다양성과 한의학 분야의 진단 과정 및 방법을 표현하는 것에 대한 모호함 등 여러 가지 이유를 들 수 있을 것이다.

하지만, 한의학 온톨로지의 계속적인 개발과 확장 등의 업데이트를 통해 한의사의 경험적 지식을 체계화하여 온톨로지에 반영함으로써 한의학 지식이 체계적으로 정리되고 분류될 수 있게 해준다.

현재 본 연구에서 모델링한 온톨로지는 비전문가에 의해 구현된 작은 형태의 온톨로지로서 한의학이라는 학문 분야에 대해 심도 있게 접근하지는 못하였지만 향후 전문가의 지속적인 도움을 받아 한의학의 진단 과정을 보다 체계적으로 모델링하고 반복적인 피드백 과정을 통하여 계속해서 업그레이드 할 예정이다.

또한, 한의학 온톨로지를 진단보조시스템에 접목시킴으로써 도메인 전문가가 아닌 비전문가(한의사)도 손쉽게 임상에서 나온 지식을 온톨로지로서 추가, 삭제, 수정함으로써 체계화되고 정리된 한의학 지식베이스의 구축이 가능해질 것이다.

이는 나아가 각각의 한의사들의 경험적 지식이 반영되어 생성된 한의학 온톨로지를 합성 및 병합을 통하여 하나의 한의학 온톨로지로서 융합하고, 융합된 한의학 온톨로지를 기반으로 전문가시스템 구현 및 한의학 지식의 체계적 공유에 도움을 줄 것이다.

참고문헌

- [1]Bossam, ETRI , <http://bossam.wordpress.com/>
- [2]박경모, 임희숙, 박종현, “Protege를 이용한 한의학의 구조화된 증상 입력을 위한 온톨로지 개발”, 대한동의생리학회, pp.1151-1156, 2003
- [3]이현실, 이두영, “온톨로지 기반 한의학 처방 지식관리 시스템 설계에 관한 연구”, 정보관리학회지, 제20권, 제1호, 2003
- [4]Available at <http://www.opengalen.org>
- [5]Available at: <http://sig.bioistr.washington.edu/projects/da/>
- [6]문경실, 박수현, 진단보조시스템을 위한 한의학 온톨로지 설계 및 구축, 2009.2
- [7]문경실, 홍승욱, 박수현, 한방 온톨로지 설계 및 구현, 2007.04
- [8]최정연, 문경실, 박수현, 한의학 온톨로지 기반의 진단 보조시스템 개발, 2008.10