

한방 온톨로지의 설계 및 구축

문경실* · 홍승욱* · 박수현**

*동서대학교 디자인&IT 전문대학원, **동서대학교 컴퓨터정보공학부

Design and Implementation of Oriental Medical Ontology

Gyung-sil Moon* · Seung-wook Hong* · Su-hyun Park**

*Dongseo University Graduate School of Design&IT, **Dongseo University Department of Computer Engineering

E-mail : siri1210@hanmail.net, hswchaos@hanmail.net, subak@dongseo.ac.kr

요 약

인터넷의 생활화와 컴퓨터의 대중화로 인해 의료 분야에서도 웹 환경에서의 지식기반 시스템의 필요성이 증대되고 있으며, 의료정보 자원과 지식기반의 응용시스템이 점차 늘어나고 있다. 또한, 지식기반 시스템에서 의료정보를 의미적 처리하기 위한 의료분야의 온톨로지 개발도 많이 시도되고 있다. 한방 온톨로지는 질병, 증상 등 정보를 체계적이고 논리적으로 표현함으로써 의료정보의 효과적 제공 및 의료지원이 가능한 지식 자원이다. 한방 온톨로지 기반의 지능화된 의료 서비스는 한방분야의 질병 및 증상 정보에 정확성을 부여하고, 체계적이고 질적으로 향상된 데이터를 제공할 수 있다. 또한, 환자의 증상에 따른 연관 증상을 제공하여 지능적인 문진이 이루어질 수 있도록 문진시스템에 적용이 가능하다. 본 논문에서는 한방 의료 분야의 질 높은 의료지원 서비스 제공을 위한 한방 온톨로지 설계 및 구현 방법을 제시하고, 구축한 한방 온톨로지를 소개한다.

키워드

온톨로지, 한방 온톨로지, OWL

1. 서 론

전통적으로 병원에서는 종이차트를 사용하여 왔으나, 인터넷이 생활화 되고 컴퓨터가 대중화됨에 따라 병원에서도 전자차트를 사용하는 횟수가 늘어나고 있다. 전자차트를 사용함으로써 얻는 효과는 우리가 상상하는 것보다도 월등히 우수하고 많은 장점을 가지고 있다. 하지만 전자차트는 아직 모든 기능면에서 완벽하지 못하고 사용되어지고 있어 종이차트에서는 기대할 수 없었던 새로운 기능들을 구현하기 위해서는 그에 따른 지식체계가 필요하다.

이는 전자차트의 단순한 정보저장과 검색 시스템에서 벗어나 지능화된 서비스가 요구되면서 시맨틱 웹(Semantic Web)을 기반으로 하는 시스템을 필요로 하는 것이다. 시맨틱 웹의 핵심적인 지식체계이자 요소가 바로 온톨로지(Ontology)이다. 의학 분야에서의 온톨로지는 의료정보의 지식표현이 가능한 지식데이터베이스라고도 한다.

한방 의료분야에서 환자 및 의사를 대상으로

질병 및 증상정보에 정확성을 부여하고 체계적이고 질적으로 향상된 데이터 제공하는 지능화된 의료지원 서비스의 제공하기 위해서는 한방 온톨로지가 필요하다. 본 논문에서는 한방분야의 질병과 증상과의 관계를 온톨로지라는 지식체계로 구축하여 한방 분야의 지식을 표현하기 위한 기본 지식체계를 제시한다.

1.1 온톨로지의 정의

온톨로지에 대한 대표적인 정의는 Gruber.T.R에 의해 정의된 'An ontology is a specification of a conceptualisation'은 온톨로지는 특정 분야에서 용어들과 그들 간의 관계를 명시적이고 정형화한 명세이다[1]. 이 정의에서는 개념화하는 방법에 따라 같은 대상에 대한 다른 표현방법의 혼란을 없애기 위해 온톨로지의 개념을 정형화시켜서 보여주는 것을 말하고 있다. 또 다른 정의로 W3C에서는 '실세계 또는 특정분야에 존재하는 모든 개념들과 그 개념들의 속성들, 그리고 그 속성들의 관계로 구성된 일중

의 사전' 이라고도 한다[2].

1.2 한방 온톨로지의 필요성

국내의 온톨로지 개발은 많은 분야에서 연구 수준에 머무르고 있다. 특히 한방분야의 온톨로지는 여러 곳에서 연구·개발 이루어지고 있으나 성과를 발표한 곳은 아직 없다. 시멘틱 웹이 대두되고 시스템의 지능화가 요구되면서 한의학 분야에서도 온톨로지 개발 필요성이 증가하고 있다.

한방 온톨로지의 개발은 온톨로지 기반의 지능화된 의료서비스를 지원함으로써, 한방 분야의 질병 및 증상 정보에 정확성을 부여하고, 체계적이고 질적으로 향상된 데이터를 제공하여, 환자에게는 자가진단의 기회 제공, 의사에게는 지능적인 문진정보 제공이 가능하게 한다.

1.3 논문의 구성

본문에서는 2.1절에서 개발 언어 및 저작도구에 대해 알아보고, 2.2절에서 의료 온톨로지의 구축 예, 2.3절에서 한방 온톨로지의 설계에 대한 살펴본다. 3.1절에서 한방 온톨로지의 구축 방법론, 3.2절에서 한방 온톨로지의 설계, 3.3절에서 한방 온톨로지의 구현, 3.4절에서 구현과정에 대해서 기술하고, 결론에서는 한방 온톨로지의 결과 및 문제점(과제), 향후 계획에 대하여 알아본다.

II. 관련 연구

2.1 개발 언어와 구현을 위한 저작도구

본 연구에서는 온톨로지 구축언어로 OWL을 선택하였다. OWL은 웹 온톨로지 관련 지식을 정의하는 언어인 동시에 웹상에서 첨단 웹 검색 및 지식관리 기능을 제공하는 온톨로지를 발간 및 공유하기 위한 시멘틱 웹 생성 언어로 DAML+OIL언어로부터 시작하였다. OWL은 단지 사람에게 정보를 표시하는데 그치지 않고 정보의 내용을 직접 처리할 수 있는 어플리케이션을 구현하는데 활용될 수 있도록 설계된 언어이다. 또한, 추론 시스템에 축적된 명제들을 정의하며, 클래스 및 그 구성원 간의 관계를 기술하고, 구문적으로 정의되지 않은 사실의 논리적 유추를 가능하게 하는 클래스 및 속성과 이에 적용할 수 있는 제약 사항의 집합으로 되어 있다[3].

본 논문에서는 온톨로지 언어의 차세대 표준인 OWL과 온톨로지 구현도구인 Protege를 이용하여 온톨로지를 구현하였다. Protege는 Stanford 대학에서 개발하였으며, Ontology를 위한 OWL 파일 작성을 위한 소프트웨어로 기획하는 웹 서비스에서 사용될 개념 등을 오브

젝트화 하고, 이들의 관계 및 속성 등을 쉽게 정의할 수 있다. 지식획득의 병목 현상을 줄이기 위해서 고안된 도구로서 1988년 경 뮤센(Musen)에 의해서 개발 되었으며, 1999년 protege-2000이 발표된 이후 현재 Protege 2006이 발표되어 많은 사람들이 사용하고 있다[4].

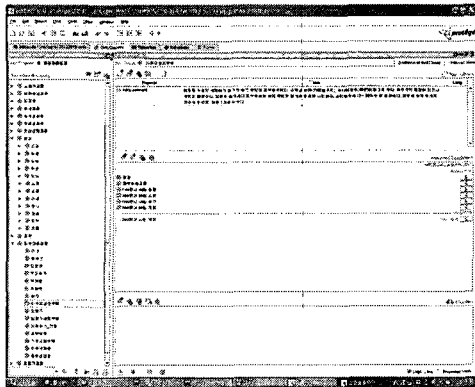


그림 1. Protege 저작도구 화면

2.2 의료 온톨로지의 구축 예

일반적인 온톨로지의 구축사례는 많지만 의학 분야, 특히 한방 온톨로지 연구는 거의 발표된 바가 없으며 현재 계속해서 진행 중인 연구 분야이다. 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[5]. ABFP ontology는 건강 상태, 증상들, 처치 등에 대한 추상적 기술을 제공하고 있으며 임의의 수만큼 인스턴스(Instance)를 생성할 수 있게 되어 있다[6]. 마지막으로 Digital Anatomist Ontology는 해부학 정보에 대한 프레임 기반의 온톨로지이다[7].

III. 설계 및 구현

3.1 한방 온톨로지의 구축 방법론

온톨로지의 개발단계에서는 명세화, 개념화, 형식화, 통합화, 구현, 유지보수 등의 과정을 거친다[8]. 명세화 단계에서는 사전검사와 같은 구축 전 단계로, 각종 자료 수집 및 지식, 정보 수집 활동이 함께 이루어진다. 개념화는 일정한 현상이나 사물이 갖고 있는 공통적인 속성을 추출해 그것들을 대표할 수 있는 포괄적인 개념을 이끌어 내는 단계이며 형식화 단계와 구현 단계를 거쳐 마지막으로 유지보수 단계에서는 수정, 정정, 추가 등 계속적으로 개선되고 변경되어 사용가능하게 해야 하기 때문에 가장

중요하면서도 어려운 단계이다. 그래서 유지보수의 문제점을 보완하게 위해 많이들 온톨로지의 자동생성, 반자동 생성에 대한 연구가 지속적으로 이루어지고 있다.

본 논문에서는 한방 온톨로지를 구축하기 위해 Protege 3.2.1 도구를 사용하였고, 한방분야의 질환(질병)을 중심으로 그에 따른 증상을 정의하고 속성들의 관계를 연결하였다.

온톨로지 구축을 위해서 질병을 최상위 클래스로 정신신경질환, 두면경부질환, 순환심장질환, 호흡기질환, 근육골격질환, 피부미용질환, 소화기질환, 비뇨기질환, 여성질환, 남성질환, 노인질환, 소아질환, 암질환, 유전질환 등을 서브클래스로 두었다. 그 아래로 세부 질병을 서브클래스로 만들었다. 이들 클래스간의 관계는 IS-A 관계를 기반으로 구현하였다.

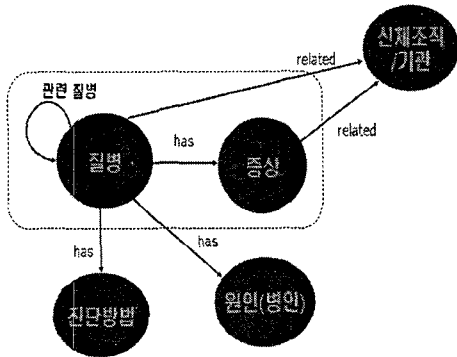


그림2. 한방 온톨로지 전체 개념 모델링

3.2 한방 온톨로지의 설계

온톨로지는 특정 용어와 속성으로 구성되어 있다. 본 논문에서는 질병 용어와 증상을 선정·분류하였고, 이들의 속성을 정의하였다. 또한, 개념을 클래스로 생성시켜, 각 개념간의 속성 관계를 부여하였고 질병에 따른 증상, 증상에 따른 질병을 구분하였다.

한의학의 병리이론에 따르면 '한의학은 단일하고 간단한 이론적 기반위에서 있는 것이 아니며, 한 다양하고 변화가 많다. 그리고 이질적인 요소들이 병존하거나 대립하고 있는 복합적인 것이다.'라고 하였다[9]. 이는 한의학이 서양 의학과와의 차이점을 보여주는 대목이 된다. 한의학에서는 질의 및 논리 추론을 통하여 진단을 내리는데 그것을 유사성(Similarity)과 유비추론(Analogical)이라고 요약할 수 있으며 이에 대한 온톨로지로서의 표현은 한의학 지식의 적절한 표현에 꼭 필요한 것이다[10].

한방 온톨로지를 구축함에 있어 여러 증상들이 각각의 질병에 중복으로 포함되어 나타날 수 있는 의미의 복잡성을 가지고 있어 관계를 정의하고 연결하는데 어려움이 있다. 아래 그림은 여러 가지 증상이 질병으로 연결되는 예

시를 보여주는 것이다. 한의학의 한 가지 질병에 대해 한마디로 규격화시켜 표현하기 힘들고 문진과 여러 단계를 거쳐 한 가지 병을 이끌어 내는 진단을 주로 하기 때문에 질병과 증상에 대한 기술이 필요하였고, 기계가 문진에 대한 구조를 어떻게 기술할 것이며 분석할 것인지에 대한 부분도 고려해야 하는 문제였다.

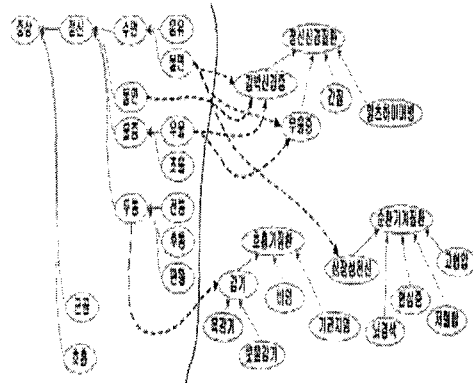


그림3. 질병-증상 연관관계 예

3.3 한방 온톨로지의 구현

3.3.1 자료의 수집과 정리

우리는 1차적으로 한의사의 도움을 받아 한방의학에서의 질병정도, 증상표, 문진표, 건강력, 진단정보입력과 관련된 자료를 수집하였고, 웹 검색을 통하여 질병분류를 추가하였다. 질환 개념 204개, 증상 개념 303개를 시작으로 만성질환 97개까지 총 301개의 질병을 개념화 하였다. 그 중에는 일반질환과 만성질환 모두에 포함되는 질병도 있다. 이와 같이 핵심 용어를 선정·분류하여 개념화를 시킨 후 계층화 작업을 하였고, 각 용어의 정의는 웹문서를 검색하여 코멘트로 추가하였다.

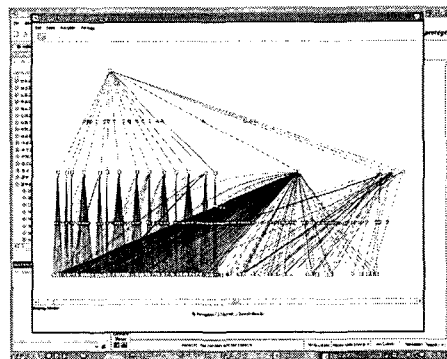


그림4. Protege 구현에 따른 구성트리

3.3.2 구현 과정

질병에 관한 핵심 용어를 선별하여 개념화시

키고 세분화하면서 개념들의 계층화를 하였다. 계층관계에서는 IS-A, PART-OF, INSTA NCE-OF 등의 관계를 적용시켜 분류하였다[11]. 한방이라는 특정 분야를 주제로 하여 핵심 용어의 개념화를 거쳐 다음과 같은 과정으로 진행하였다. 먼저, 질병과 증상을 최상위 클래스로 생성하였고, 그 아래로 세부 질병명과 증상명을 분류하였다. 세부 질병에서 나타날 수 있는 증상들을 골라 각각의 질병과 관계를 지었다. 예외적으로 한 종류의 질병이 일반 증상뿐만 아니라 또 다른 질병을 증상으로 가지고 있는 경우도 있다.

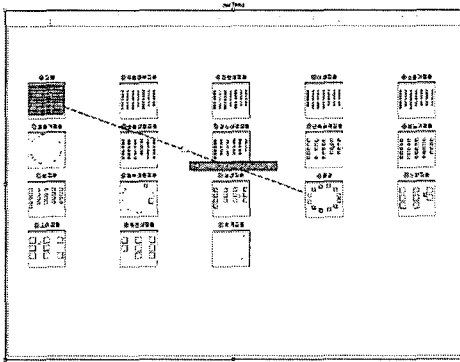


그림 5. Protege에서의 질병-증상 관계도

기본적으로 각 신체부위별 모든 질병은 그에 대한 증상, 원인, 처방을 가지고 있으며, 그들의 관계를 어떻게 정의, 연결하는가에 따라 한방 온톨로지의 완성도를 높일 수 있다. 이것의 의미는 원하는 정보를 선택하고 컴퓨터가 온톨로지라는 지식체계를 이용하여 그 정보를 찾을 때 용어의 선택에 따라 다른 답변이 나올 수 있다는 것이다. 그래서, 본 연구실에서는 전문 용어와 유사용어의 분류와 표현에 많은 노력을 기울였다.

IV. 결 론

정보의 바다에서 웹 문서는 엄청나게 쏟아지고 있다. 그 가운데에서 내가 원하는 정확한 정보를 찾기란 어렵다. 원하는 정보를 고르기 위해 수천, 수만 가지의 웹 문서를 열어봐야 하고, 정확하지 않는 정보를 수용하게 된다. 이러한 문제점을 해결하기 위해 우리는 온톨로지를 구축하고, 그것을 이용함으로써 보다 정확하고 체계적인 정보를 얻을 수 있다.

한방 온톨로지를 구축은 많은 부분에서 어려움이 있다. 수작업으로 시작된 한방 온톨로지의 구축은 분명히 한계가 있다. 수동적인 온톨로지 구축을 자동화 시키므로, 대량의 온톨로지

를 쉽고 편리하게 구축하여 효율적인 검색이 가능하도록 온톨로지의 반자동생성, 자동생성에 대한 연구도 추가적으로 이루어져야 한다.

참고문헌

- [1] Gruber.T.R., A Translation Approach to Portable Ontology Specifications, Knowledge Acquisition, 5(2), pp. 199-220.1993
- [2] Protege, <http://protege.stanford.edu>,2006.7 A Research on web Ontology Development Guideline)
- [3] OWL Web Ontology Language Gride, <http://www.w3c.org/TR/owl-guide>,2004.10
- [4] protege, <http://protege.stanford.deu>.2006.7
- J. Gennari, M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Cruby, H. Eriksson, N.F.Noy, S.W.Tu The Evolution of Protege An Environment for Knowledge-Based Systems Development. 2002 [Technical Report] availableat http://www.smi.stanford.edu/pubs/SMI_Reports/SMI-2002-0943.pdf.
- N.F.Noy, R.W.Fergerson, & M.A.Musen. The knowledge model of Protege-2000: Combining interoperability and flexibility. 2th International Conference on Knowledge Engineering and Knowledge Management (EKAW'2000), Juan-les-Pins, France, . 2000.
- [5] Available at <http://www.opengalen.org>
- [6] 박경모, 임희숙, 박종현 "Protege를 이용한 한의학의 구조화된 증상 입력을 위한 온톨로지 개발"
- [7] Available at: <http://sig.biostr.washington.edu/projects/da/>
- [8] 한국전산원 "웹 온톨로지 개발지침 연구"
- [9] 여인석, "한의학의 병리이론"
- [10] 박경모 "한의학 및 보완대체의학에서의 의학 온톨로지 개발" 2003.12
- [11] 한국과학기술원 최기선·류범모 "온톨로지의 구축과 학습: 상하위 관계"