

Protege를 이용한 한의학의 구조화된 증상 입력을 위한 온톨로지 개발

박경모* · 임희숙 · 박종현¹

경희대학교 전자정보대학 동서의료공학과, 1:대구한의대학교 한의과대학 병리학교실

Building an Ontology for Structured Data Entry of Signs and Symptoms in Oriental Medicine

Kyung Mo Park*, Hee Sook Lim, Jong Hyun Park¹

Department of Biomedical Engineering, Kyunghee University,

1:Department of Pathology, College of Oriental Medicine, Daeguhaany University

To obtain both of the fast and complete data entry and the acquisition of reusable data in a Computer-based Patient Record system (CPR), we are building the ontology that is used by the entry supporting agents. Our application domain is Traditional Chinese Medicine. As the tool for the implementation, we used protege 2000 which is ontology building tool and provides frame knowledge representation language. In this paper, the construction methodology of our ontology is reported.

Key words : Ontology, Structured Data Entry, Oriental Medicine, protege, Computer-based Patient Record System, Knowledge Representation

서 론

1. 의학 영역에서의 계산 가능한 지식체계의 필요성

전자진료부는 전통적으로 사용되어오던 종이 진료부에 비하여 데이터의 관리 측면에서 볼 때는 비교할 수 없을 정도의 장점을 가지고 있다. 따라서, 처방전달시스템과 같은 전자진료부 시스템의 일부 요소들은 이미 성공적으로 개발되어 많은 비용을 줄이고 있다. 그러나, 전자진료부가 환자 정보의 '의미를 다루어야만 수행할 수 있는 일'에 관해서는 아직도 극히 일부만이 수행되고 있을 뿐이다. 이렇게 의미를 다루어야 수행할 수 있는 일의 범주에는 환자 정보의 검색, 환자 정보의 통합적 분석, 진료에 관한 의사결정의 보조 등 전통적인 진료부에서는 기대할 수 없었던 새로운 기능들이며, 앞으로 구현되어야 하는 전자진료부의 핵심기능들이다. 이러한 기능들이 수행되려면, 전자진료부 시스템이 사용할 수 있는 형태로 표현된 지식체계가 필요하며, 그러한

지식체계를 다룰 수 있는 프로세스(Computational Process)가 제안되어야만 한다. 이와 같은 기계가 해석 가능한¹⁾ 지식체계를 본 논문에서는 온톨로지(Ontology)라고 부르며, 전자진료부가 원하는 지식을 온톨로지를 이용하여 제공하는 프로세스의 총합을 에이전트(Agent)라고 부른다.

2. 온톨로지(Ontology)의 도입

온톨로지는 원래 철학의 한 분과로서 '존재들에 대한 설명들'이란 의미의 존재론을 말한다. 정보과학에서 온톨로지는 특정 영역의 용어들과 그들간의 관계를 명시적이고 정형화한 명세(Explicit formal specifications of the terms in the domain and relations among them)로 정의된다.^[1]

종이진료부를 작성하는 한의사에게 한의학의 지식, 진료부 작성 요령, 진료부 표기에 대한 규약 등이 필요한 것처럼 지능적인 전자진료부를 구현하기 위해서는 온톨로지가 필요하다. 온톨

* 교신저자 : 박경모, 경기도 용인시 기흥읍, 경희대학교 동서의료공학과
· E-mail : saenim@khu.ac.kr · Tel : 031-201-2979
· 접수 : 2003/06/24 · 수정 : 2003/07/30 · 채택 : 2003/09/29

1) 여기에서 해석이라는 용어는 인간의 '인지과정'을 그대로 재현한다는 의미에서 사용한 것이 아니고, 인간의 '인지 과정'이 보이는 결과를 그대로 흉내낸다는 의미에서 사용한 것이다.

로지는 넓은 의미에서 환자정보를 포함한 의무기록을 기술하기 위해 필요한 지식을 기록하고 있는 데이터베이스라고 간주할 수 있으며, 지식과 관련된 데이터라는 의미에서는 지식베이스(Knowledge-base)이다. 그러나, 온톨로지는 그 자신이 담고자 하는 영역의 지식을 좀 더 그대로 표현하기 위해서 표현력을 높인 것으로, 따라서 전통적인 데이터베이스에 비하여 좀 더 표현력이 큰 논리적인 언어에 의해 표현되고 있다. 온톨로지 연구는 원래 인공지능의 지식 표현(Knowledge Representation) 방법에서 출발하였으며, 온라인의 지식처리 기능을 요구하기 시작하면서 Semantic Web 연구를 계기로 활성화되고 있다.

본 연구에서 개발한 프레임 기반의 온톨로지도 W3C(World Wide Web Consortium)에서 제안하는 RDF schema에 의해서 기술(記述)되었다. 이는 향후 온라인에서의 응용과 상호 호환 가능한 지식체계를 구축하기 위한 전략이다.

3. 기존의 의학 영역에서의 온톨로지들

현재까지 개발된 진료 영역에서의 온톨로지 중 주목할 만한 것은 GALEN, Digital Anatomist Ontology, ABFP(American Board of Family Practice) Ontology 등이 있다. GALEN은 Description logic 기반의 GRAIL(GALEN Representation and Integration Language)에 의해서 표현된 온톨로지로서 맨체스터 대학을 중심으로 하는 유럽 연합에서 개발되었다.^[2] Digital Anatomist Ontology는 해부학 정보에 대한 프레임 기반의 온톨로지이다.^[3] ABFP ontology는 건강 상태, 증상들, 처치 등에 대한 추상적 기술을 제공하고 있으며 임의의 수만큼 인스턴스(Instance)를 생성할 수 있게 되어 있다. 이 밖에도 가이드라인 기반 진료(Guideline-based care)를 위한 온톨로지와 그와 관련된 여러 프로젝트들이 수행되고 있다. 기존의 용어시스템으로 개발된 것들 중에서도 온톨로지의 특성을 가진 용어체계로는 UMLS나 SNOMED CT²⁾를 거론할 수 있다.^[4,5] 그러나, 기존의 용어시스템과는 그 용도에서 차이가 있으며 UMLS나 SNOMED를 포함하여 기존의 용어 시스템은 아직까지 지식표현을 위한 언어들이 가지고 있는 속성(property)의 이행성(transitivity), 카디널리티(Cardinality), 인스턴스의 표현, 속성의 상속(Inheritance) 등을 비롯한 개념들을 처리하기 위한 논리적 추론 기능을 가지고 있지 못하다.^[6]

한의학에서의 필요성과 요구

1. 컴퓨터를 통한 한의학 지식의 처리 기술

환자의 증상은 단순히 규격화된 증상명만으로 표현해서는 안 되는 경우가 많다. 즉, 좀 더 상세한 기술이 필요하고 그래야 만이 비로소 변증을 위한 가치를 가지게 된다. 그러나, 이러한 상세한 기술은 다른 난점을 가지게 된다. 우선, 기술(記述)하기가 더 어려우며, 단순히 텍스트로서 기술되는 한 임상 연구 등과 같

은 곳에 재사용 될 수 없으며, 전산화시켜 계산 가능한 형태를 만들어 내지도 못한다. 따라서 증상을 상세하게 기술하되 좀 더 유용한 정보의 형태를 가지도록 할 필요가 있다. 그러나 환자가 호소하는 복잡한 증상을 단순한 증상명이나 병명만으로 표현하는 것은 정보량의 부적절한 축소일 뿐 아니라, 실제로 진단을 위해서도 충분한 자료가 되지 못한다. 또한 한의학에서는 증상자체 뿐 아니라 그 증상이 가지고 있는 속성이 진단에 있어서는 더 필요한 정보를 제공하기 때문에 증상의 특성에 대한 기술이 필요하다. 예를 들어 '두통'의 경우에도 두통의 부위, 통증의 속성, 동반하는 증상, 유발 원인 등의 속성들이 변증을 위해서 필수적인 요건이 된다. 동일한 이유로, 임상 연구 시에 체계화되고 정리된 증상의 기술(記述)이 반드시 필요하다. 한의학은 증상을 포함하여 정량화하기 힘들고 단위화하기 힘든 문진 정보들을 통해서 진단을 하고 있기 때문에 이에 대한 대안이 필요하다. 통계적 분석의 예를 볼 때도 복잡한 증상 기술들의 기본적인 개념으로의 단위화와 함께 그들 간의 관계를 기술하는 것이 필요하다. 이는 곧 증상을 자세하게 기술하되 이들이 기계적으로 분석 가능해야만 하는 것을 의미한다. 또한 앞에서 기술하였다시피, 한의학 분야에서도 다른 의학 분야와 마찬가지로 전자진료부 시스템, 문서 검색 시스템, 학술문헌 관리 등에서 새로운 기능들을 요구하고 있다.

이상의 요구에 부응하는 새로운 기술이 한의학의 지식을 표준화된 방법들로 기술(記述)하며 의미 기반으로 다룰 수 있는 온톨로지 기술(技術)이며, 이것을 통해서 지식표현이 가능할 수 있게 된다. 그런데 현재는 한의학과 관련된 온톨로지는 전 세계적으로 아직 개발된 바가 없으며,³⁾ 가능한 빠른 시일 내에 지식공학 분야에서의 요구를 반영하기 위해서는 한의학의 온톨로지를 개발해야만 한다.

2. 구조화된 데이터 입력을 위한 온톨로지의 필요

본 연구에서 개발하고자 하는 온톨로지는 앞에서 기술한 전자진료부의 구조화된 증상 입력을 위해서 고안되었다.^[13](Fig. 1)

본 연구팀이 고안한 전자진료부의 인터페이스를 보면 Fig. 1에서 보는 바와 같다. 증상 항목을 선택하면 작업창 부분에 template이 자동으로 생성된다. 생성된 template의 내부에 기술될 concept과 relation은 프로그램 우측에 있는 메뉴 항목을 선택하면 된다. 이들의 관계를 설정하는 것은 CGs에서 사용하는 relation과 arc 개념을 도입하여 표현하도록 하였다.

template을 통해 기술된 하나의 증상 object는 다른 증상 object와도 관계를 맺게 된다. 또한 새롭게 만드는 증상 object 뿐 아니라, 해당 환자 domain에서 예전에 기술된 내용을 다시 불러와서 사용할 수도 있고, 환자가 진료를 받기 전에 먼저 입력된 필수 입력 항목이나 과거력, 가족력 등의 내용도 하나의 object로 사용, 관계를 맺을 수 있다. 이러한 증상 object들간의 관계를 설정함으로써 의사는 환자와의 대화를 통해 기록해야 될 부분을 모두 컴퓨터에 입력하게 되는 것이다. 이렇게 자동적으로 임의의

2) De Keizer et al(2000)은 SNOMED RT가 온톨로지가 아니라고 분석하였지만, 2002년과 2003년에 걸쳐 개정된 SNOMED CT는 개념들간의 관계(relationship)들이 기술되어 있으며 제한된 형태의 프레임 로직을 가진 온톨로지라고 말할 수 있다.

3) 국내의적으로 설계 단계에 있는 한의학 관련 온톨로지들에 대해서 직간접적으로 접한 바는 있지만 아직까지 공식적으로 제안되거나 유포된 온톨로지는 보고 받은 바가 없다.

증상을 기술할 수 있게 하기 위해서는 전자진료부가 한의학의 관련 지식들을 보유하고 있어야 하며, 본 연구팀이 그것을 구현하기 위해서 만든 것이 본 논문에서 기술하는 온톨로지이다.

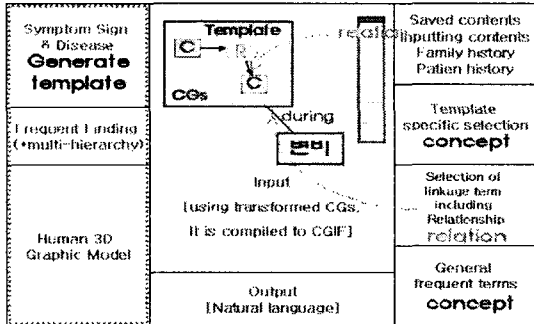


Fig. 1. program interface schema

본 연구팀은 전자진료를 위한 솔루션의 하나로써, 구조화된 증상 입력 에이전트를 개발하고 있으며 그 에이전트가 사용할 지식베이스로 온톨로지를 설계하고 그 일부를 구축하였다. 온톨로지 개발 영역으로는 한의학적인 특징을 담고 있으면서 임상영역의 정보를 표현하는데 가장 필요한 부분인 증상(Sign & Symptom) 영역의 일부를 채택하였다. 이에, 한의학 분야에서의 온톨로지를 개발하기 위해서 필요한 전반적인 방법론을 보고하고자 한다.

온톨로지 개발 방법

1. 온톨로지 개발 언어와 도구

본 연구팀은 온톨로지 개발 언어로 지식표현 분야에서 전통적으로 사용되던 프레임 기반 지식표현 언어를 선택하였다. 프레임(frame)은 1975년 민스키(Minsky)에 의해 제안되었으며, 그것의 기본 구조는 기술대상이 되는 프레임(frame)과 프레임의 속성을 표시하는 슬롯(slot), 그리고 그 슬롯에 들어가는 속성 값인 파셋(facet)으로 구성된다. 이들은 최근에 W3C에서 발표된 RDF와 RDFs에 의해서 표현 가능하다. 또한, 온톨로지의 구축을 위한 도구로서는 protege-2000을 사용하였다. protege는 지식획득의 병목 현상을 줄이기 위해서 고안된 도구로서 1988년 경 뮤센(Musen)에 의해서 개발되었으며, 1999년 protege-2000이 발표되었다.^[7,8] protege는 전문가 시스템 그 자체도 아니고 전문가 시스템을 제작하기 위한 프로그램도 아니다. 어떤 특정한 응용영역에서의 전문가 시스템을 만들기 위해서는 지식 획득이 요구되는데, 이러한 지식 획득을 지원하기 위해서 주문제자 되어지는 도구들을 만드는데 도움을 주는 도구가 protege이다.^[9]

이러한 프레임 기반의 지식표현은 논리적 단순성과 사용의 편의성, 계산적 복잡성 면에서 이점이 있지만 표현력이 현저하게 떨어지기 때문에 프레임 기반의 온톨로지는 응용프로그램의 요구를 충분히 수용하지 못한다. 따라서, 본 연구에서는 프레임 시스템을 지원하는 protege를 사용되 그에 대한 보충으로 전통적인 프레임 기반 표현 방법을 사용하지 않고 수정 사용하였다.

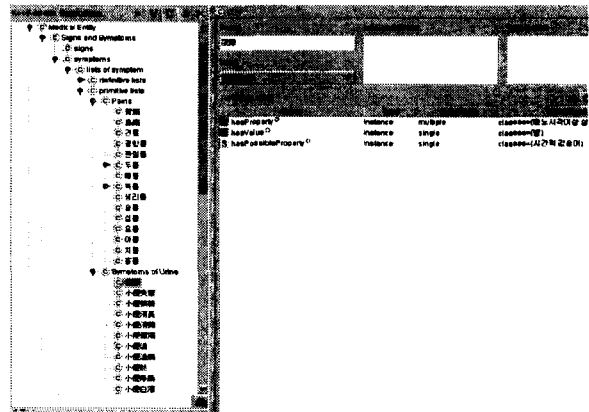


Fig. 2. Screen shot of protege-2000

2. 한의학 증상의 수집과 정리

우리가 한의학의 문헌을 1차 자료로 사용하거나 또는 문헌을 분석하여 재정리하여 일차진료를 위한 증상 개념을 약 600개 얻었다. 사용할 개념의 수는 온톨로지의 개발 방향, 실제 응용분야의 특성들에 의해서 달라질 수가 있다. 왜냐하면 온톨로지에 이미 표현된 개념들과 그들 간의 관계는 온톨로지에 표현되지 않은 지식의 표현을 위한 토대로 사용될 수 있기 때문이다. 즉, 이론적으로 지식의 토대론(fundamentalism) 입장에서 볼 때 공리에 기반한 온톨로지(axiomatized ontology)는 지식의 기반이 되는 개념만 온톨로지에 기술되면 그 기반으로부터 파생하는 모든 개념들은 온톨로지에 등록되어 있지 않아도 얼마든지 표현될 수 있다. 즉, 온톨로지의 언어가 1차 술어 논리의 표현력을 가지고 있으며 기본적인 원시 개념(Primitive concept)이 충분하다면 이론적으로는 일정한 범위 내에서는 어떠한 수준에서도 온톨로지의 개발 범위를 결정할 수 있을 것이다. 그러나, 실제로는 우리가 채택한 프레임 기반의 표현 방법이 가진 논리적 한계와 우리가 구축한 온톨로지를 기반으로 개발될 실제 응용프로그램의 특성을 고려하여 표현할 개념의 수를 정하였다.

개발된 온톨로지의 개요

1. 도메인(Domain)의 영역의 구조

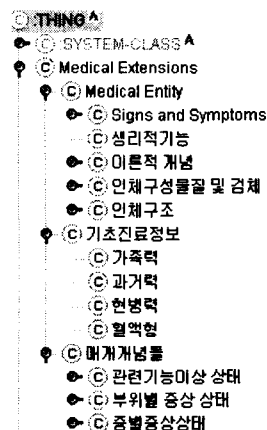


Fig. 3. Hierarchy of medical extension in our ontology

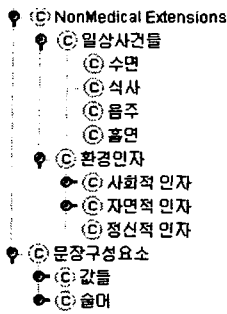


Fig. 4. Hierarchy of non-medical extension in our ontology

현재 구현된 온톨로지는 상위 수준의 온톨로지를 가지고 있지 않은 채 본 연구팀의 요구에 부합하는 부분만이 구축되어 있다. 즉, 한의학의 지식구조를 조망할 수 있는 추상적이고 철학적인 상위 수준의 존재들에 대한 토의는 아직 이루어지고 있지 않다. 다만, (Fig 3, Fig 4.)와 같이 증상과 관련된 한의학 영역의 개념들과 그와 관련된 비의학적 영역의 개념들의 일부를 대상으로 하였다. Fig 2,3은 본 연구팀이 구축한 온톨로지의 계층구조를 보여주고 있다. 이 계층구조는 각 개념 같이 모든 관계들 중 일부를 보여 주는 것으로 Subsumption의 관계이다.

2. 지식표현 방법론 : 온톨로지의 구성요소들

현재 우리는 환자 정보의 입력을 위한 온톨로지를 네 개의 구성부분으로 구분하여 구축하고 있다.(Fig. 5.)

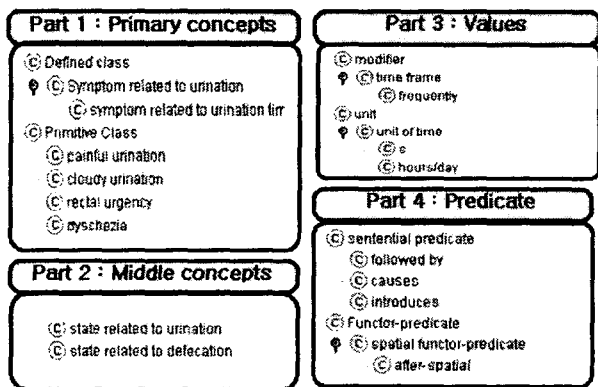


Fig. 5. Components of our ontology

첫 번째 구성물은 '1차적 개념'으로서 그 속성이 기술되어야만 하는 의학적 개념들로 구성되어 있으며, 두 번째 구성물로서 우리는 프레임 기반의 지식표현이 가지는 문제를 보완하기 위한 장치로서 '매개개념'을 도입하였으며, 이는 Description logic을 기반으로 하는 온톨로지의 구축에서 온톨로지를 정규화시키는 방법론으로부터 착안된 방법이다.^[10] 세 번째 구성물은 1차적 개념의 속성을 기술하기 위한 '값 타입'으로 구성되어 있다. 마지막의 '술어'는 프레임 기반의 지식표현이 표현하지 못하는 부분의 표현을 위해서 만들어진 술어 기반의 표현을 위한 개념과 장치들로 구성되어 있다. 이에 대한 상세한 설명은 아래의 절에서 설명된다.

3. 지식표현 방법론 : 고유속성과 부가속성의 도입

본 온톨로지의 1차적 개념의 입장에서 볼 때, 증상을 기술하는 방법은 다음의 두 가지로 요약된다. 하나는 증상이 본질적으로 고유하게 가지고 있는 필연적 속성의 표현 방법이고, 다른 하나는 증상이 가질 수 있는 우연적 속성의 표현방법이다. 온톨로지에서는 전자만을 표현하고, 후자는 전자진료부에서 사용자가 표현할 수 있게끔 그 방법을 온톨로지에서 제공한다. 본 연구팀은 전자를 고유속성(intrinsic property), 후자를 부가속성(extrinsic property)라고 명명하였다.(Fig 6, Fig 7.)

본 온톨로지를 통한 환자의 증상의 기술은 이미 제공되고 있는 증상 컨셉을 단순히 선택하거나 더 자세히 기술하여 증상 event를 기술하는 intrinsic pathway와 온톨로지의 개념들을 조합하여 새로운 증상 사건을 기술하는 extrinsic pathway로 나누어지며 이들은 각각 고유속성, 부가속성과 관련된다. 그리고 각각의 pathway가 작동하기 위해서 온톨로지 상의 class들 간에 relation을 일정한 원칙에 의해서 구성하였다.

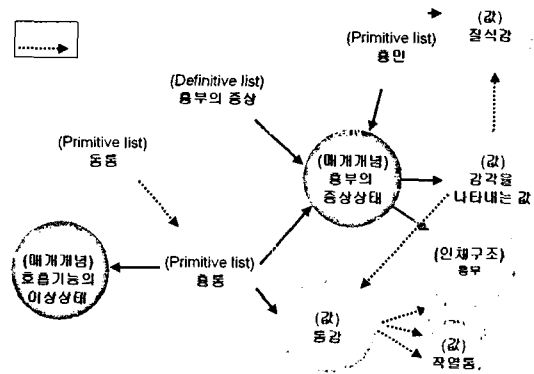


Fig. 6. The representation of intrinsic properties

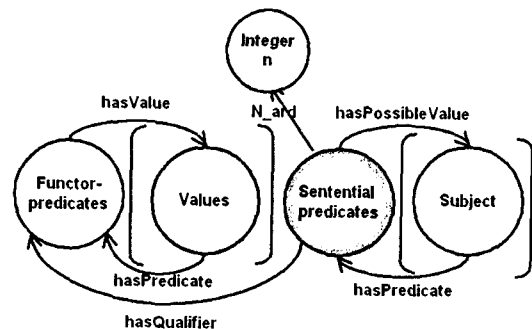


Fig. 7. The framework for the extrinsic pathway

온톨로지 구축 방법론 토의

1. 매개개념(Middle Concept)의 도입

Fig. 6의 그림에서 '(매개개념)흉부의 증상'이나 Fig 8의 '(매개개념)배뇨시간 이상상태'가 매개개념의 예이다.

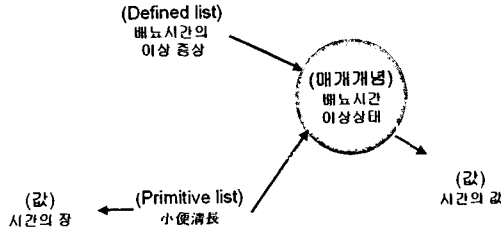


Fig. 8. An example of 'Middle concept'

이러한 매개개념의 도입은 온톨로지의 구축을 더 번거롭게 만들고, 일반적인 영역전문가[domain expert]에게 직관적으로 다가오지 않는다. 하지만, 최소한 두 가지 이상의 이점 때문에 도입하였다.

첫째는, 그렇지 않으면 설정해야 하는 multi hierarchy의 문제점을 피하기 위해서다. 이는 사용자들이 사용하는 개념들의 다면성과 중복성으로부터 기인하는 복잡함을 피하고, 내부에 단일하고 배타적인 속성을 가진 관념적인 개념을 도입하여 개념의 단위화를 시도하는 것이다. 예를 들어 매개개념을 통해 조작하지 않으면, 배뇨시간 이상 증상과 소변칭장은 'Is a' 관계 즉, subsumption 관계를 가지게 되는데 이러한 관계 맺음은 배뇨시간 이상 증상 중 하나인 소변빈삭이 가질 여러 subsumption 관계와 구별해주질 못한다. 물론, 소변칭장이 가지는 모든 상속된 속성은 당연히 소변칭장의 속성이다. 그러나, 우리는 우리의 응용 범위에 따라 일부 속성과 선별할 필요를 느끼게 된다. 그런데 이 방법은 '소변칭장'을 사용할 때 그가 상속받는 여러 속성을 적절하게 선택할 방법이 없어지게 된다. 그러나, 매개개념에 의해서 연결하게 되면 소변칭장과 매개개념 사이의 관계를 적절하게 설정할[여기에서는 hasProperty]에 따라 상속[정확히 말하면 상속이 아니고 연결]되는 개념을 사용할지 말지를 결정할 수 있게 하고자 함이다.

둘째는, 첫 번째의 문제를 해결하기 위해서 다른 시도를 할 수 있다. 즉, 배뇨시간의 이상 증상은 시간의 값을 가지고, 소변칭장은 '시간의 長' 값을 가지게 되면 자연스럽게 hierarchy가 생성될 수 있다. 그러나, 그것은 현재로는 다른 defined class[예를 들어 배뇨시간의 이상 증상]가 동일하게 시간의 값을 속성으로 가질 때 그 들간의 차이점을 구분 할 수 없게 된다. 이것을 해결하기 위해서는 값의 종류를 더 다양하게 만들어야 하는데[예를 들면, 배뇨 시간의 값과 배변시간의 값 등과 값이] 이것은 매개개념을 만들어 내는 것과 똑같은 복잡도를 야기하고, 더 나아가 다른 문제를 유발한다. 추가적으로 multiple hierarchy는 계산의 복잡성을 유발한다

2. 맥락(context)의 표현

하나의 개념은 어떤 하나의 객체나 사건을 의미한다. 그러나 실제로는 어떤 하나의 객체가 여러 독립적인 의미들을 함축하고 있는 경우가 많다. 인간의 경우에는 무의식적으로 다른 맥락에서

다른 의미를 선택하여 사용할 수 있다. 그러나, 현재에 구축된 많은 온톨로지들은 그러한 맥락을 표현하지 못하고 여러 계층 구조상에 하나의 개념을 위치하게 하고 있다. 이러한 방식으로는 맥락적 독립성을 표현할 수 없다.

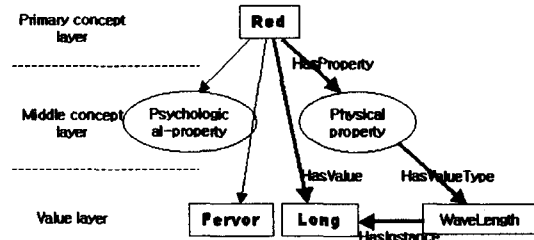


Fig. 9. Framework of middle concept

본 연구팀은 프레임 로직에서 사용하는 triple[class, slot, filler] 구조를 변형하여 [class, relationship, class]와 같은 형태로 사용함에 따라 relation을 이용해서 개념들간의 관계를 표시하고 있다. 이러한 상태에서 개념 정의 문장(Concept Defining sentence;CD)을 맥락 지시 문장(Context Inducing sentence;CI)과 연결하는 매개개념을 도입하였다.

Fig 9에서 와 같이 [Red, hasValue, Long]와 [Red, hasProperty, PhysicalProperty]는 각각 개념 정의 문장(CD)과 맥락 지시 문장(CI)이며, [PhysicalProperty, hasValueType, WaveLength]와 [WaveLength, hasInstance, Long]는 모든 CD-CI의 쌍에서 고정되어 있는 문장이다.

3. 유사성의 표현

유비 추론을 구현하기 위해서는 어떠한 개념들이 서로 다른 맥락에서는 주된 개념(Primary concept)을 중심으로 다시 재배열 되어야만 한다. 매개개념은 분류 범주로 작동하는 값(value)들을 이용하여 주된 개념들을 재배열할 수 있다.(Fig 10)

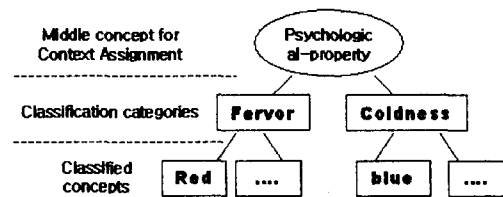


Fig. 10. Autonomic rearrangement by context

4. 계산적 복잡도의 축소

실제 온톨로지에서 개념(Class)의 수(Nc)는 관계(relation)의 수(Nr)보다 상당히 많기 마련이다. 특정한 맥락을 표현하기 위해서는 본 연구팀이 고안한 온톨로지 방법론은 Nr를 확장하기보다는 Nc를 확장하는 방법을 택하였다. 표현력이 같은 온톨로지들에 있어서 Nc와 Nr의 총계가 상수 K로 고정되어 있을 때 계산 복잡도(Computational complexity)의 중간값(Ce)은 Nc가 커질수록 줄어든다. 왜냐하면 실제 Nc와 Nr의 비율은 Fig 11에서 오른쪽 굵은 선 부분이기 때문이다.

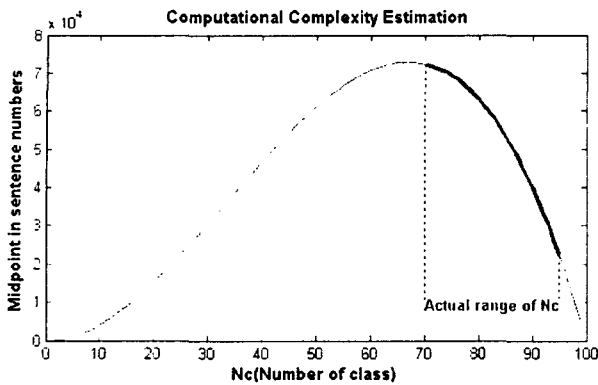


Fig. 11. Simulation of computational complexity

$$Nc + Nr = K \quad (1)$$

$$s = Nc * (Nc - 1) * Nr \quad (2)$$

$$Ce = \frac{\sum_{a=0}^{s-2} (n-1+a) \cdot s-2 C_{a+1}}{\sum_{a=0}^{s-2} s-2 C_{a+1}} \quad (3)$$

전자진료부와 같은 응용 프로그램 개발자의 인지적인 측면에서도, 개념이나 관계 중의 어느 하나의 비율이 압도적으로 많은 경우에 부담이 줄 것으로 예상된다.

결 론

본 연구팀의 온톨로지는 한의학을 위한 전자진료부 시스템의 다른 모듈, 진단을 위한 추론, 데이터의 분석, 자연어 처리 등을 위한 온톨로지로서 사용할 수 있도록 범용성을 고려하여 설계되었으며, 프레임 기반의 Protege 2000을 사용하여 한의학의 증상 온톨로지를 구축하였고 그것을 RDFs 포맷으로 exporting시켜서 응용프로그램을 위한 에이전트를 구현하고자 하였다.

온톨로지 표현언어를 선택하는 문제에 있어서 충분한 표현력(Expressiveness)과 적절한 계산적 복잡성(Computational Complexity), 관련분야의 표준과의 호환성 등을 고려하면 Description logic을 기반으로 하는 2003년 초에 발표된 OWL과 그 이전에 발표된 DAML+OIL이 가장 적절한 것으로 생각된다.[11] 그러나, 아직 여러 기술적인 지원이 충족되어 있지 않기 때문에 본 연구에서는 사용하지 못하였다. 또한, 한의학의 진단을 위한 추론은 전형적인 유비추론(analogical reasoning)을 이용하고 있기 때문에 이를 위한 온톨로지의 구축은 상당히 흥미로운 도전이 될 것으로 생각한다.

감사의 글

“본 연구는 보건복지부 보건의료기술진흥사업의 지원에 의하여 이루어진 것임(01-PJ1-PG3-51300-0020)”

참고문헌

1. Gruber, T.R., A translation approach to Portable Ontology specification. Knowledge Acquisition, 1993;5:199-220
2. Available at <http://www.opengalen.org/>
3. Available at: <http://sig.biostr.washington.edu/projects/da/>
4. De Keizer N, Abu-Hanna A, Zwetsloot-Schonk J. Understanding Terminology System I: Terminology and Typology. Methods of Information in Medicine 39:16-21, 2000.
5. De Keizer N, Abu-Hanna A. Understanding Terminology System II: Experience with conceptual and formal representation of structure. Methods of Information in Medicine 39:22-29, 2000.
6. D. E. Oliver, Y. Shahar, M. A. Musen, & E. H. Shortliffe. Representation of Change in Controlled Medical Terminologies. Artificial Intelligence in Medicine 15(1):53-76, 1999.
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] available at: http://www.smi.stanford.edu/pubs/SMI_Reports/SMI-2002-0943.pdf
8. 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.
9. Musen, M.A. Automated Generation of Model-Based Knowledge-acquisition Tools, Pitman publishing, 1989 http://www-smi.stanford.edu/pubs/SMI_Reports/SMI-2000-0830.pdf
10. Rector, A.L., Wroe C., Rogers J.E., Roberts A. Untangling taxonomies and relationships: personal and practical problems in Loosely coupled development of large ontologies. K-CAP'01, Oct 22-23, 2001
11. Sean Bechhofer, Ian Horrocks, Carole Goble, Robert Stevens. OilEd: a Reason-able Ontology Editor for the Semantic Web. Proceedings of KI2001, Joint German/Austrian conference on Artificial Intelligence, September 19-21, Vienna. Springer-Verlag LNAI Vol. 2174, pp 396-408. 2001. Available at: <http://potato.cs.man.ac.uk/papers/ki2001.pdf>
12. 박경모, 박종현. 한의학의 증상표현을 위한 방법론. 동의생리병리학회지 16(5):845-850, 2002
13. 임희숙, 박경모, 전자 의무기록 시스템에서 구조화된 데이터 입력을 위한 인터페이스 설계 및 구현. 한의진단학회지, 기고중