대한의료정보학회지
제9권, 제3호 Volume9, Number3

종 설

의료정보학에서의 온톨로지 기술

김홍기, 김명기[1]

단국대학교 경영정보학, 서울대학교 치과경영정보학[1]

Ontology Technology in Medical Informatics

Hong-Gee Kim, Myeng-Ki Kim[1]

Department of Management Information Systems, Dankook University
Department of Dental Management & Informatics, Seoul National University[1]

Abstract

The main purpose of this paper is to overview the current issues in the area of medical ontology. Ontology technology in Medical Informatics is evolved from the three different research areas: namely, web application for the Semantic Web, Knowledge Representation in Artificial Intelligence, and medical terminology system. In this paper we provide possible research agenda concerning medical ontology development from the above three perspectives at the same time. ***(Journal of Korean Society of Medical Informatics 9-3,213~219, 2003)***

Keyword : Ontology, Knowledge Representation, Medical Informatics, Artificial Intelligence, Semantic Web

교신저자: 김홍기, 충남 천안시 안서동 산 29번지 단국대학교 경영정보학 (330-714)
전화: 041)550-3363 Fax: 041)550-3354 E-mail: hgkim@dku.edu
* 본 논문은 " 2002학년도 단국대학교 대학연구비의 지원 "에 의하여 연구되었음.

I. 서 론

정보기술의 적용대상이 되는 여러 학문 및 산업분야 중에서 의료분야는 정보의 양과 복잡성 면에서 가장 높은 수준의 정보관리 기술이 필요하다. 매 순간마다 새로운 의학지식은 만들어지고 엄청난 양의 환자에 대한 의무기록이 생겨나고 있지만, 의료분야에서의 정보와 지식은 종이에 기록된 구조화 되어있지 않은 형태의 것들이 대부분이다. 최근 전 산업분야에 걸쳐 분산적 웹 환경에서의 효율적 정보관리를 위한 기술 개발이 활발히 이루어지고 있다. 특히 컴퓨터가 정보의 의미를 인식하고 처리할 수 있는 "시맨틱 웹(The Semantic Web)"이 차세대 웹 기술로 부각되고 있다. 의료분야도 예외는 아니다. 시맨틱 웹을 구현하는 핵심기술인 온톨로지(ontology) 연구의 수많은 성과들이 의료분야를 중심으로 이루어지고 있다. 본 논고의 목적은 의료정보학에서의 온톨로지 기술을 조망함으로써 이 분야 연구의 중요한 현안들에 대한 이해를 돕고 지금까지의 의료정보학의 연구 성과들을 온톨로지기술과 연결할 수 있는 가능성에 대하여 고찰하는 것이다.

의료정보학의 시작은 Ledley와 Lusted가 1959년에 Science지에 "Reasoning Foundations of Medical Diagnosis"라는 논문을 발표하면서라고 할 수 있다 (Ledley & Lusted, 1959). 이들은 의사의 진단과정을 컴퓨터가 처리할 수 있는 형태의 체계적인 추론절차로서 표현하는 방법을 소개하였다. 즉, 의학지식을 질병과 증상과의 관계로 표현하고 이를 천공카드에 입력함으로써 환자의 증상과 관련 있을 수 있는 질병들을 자동적으로 찾아주는 기술을 개발하였다. 그 후 본격적인 의료 진단을 위한 전문가시스템인 MYCIN이 개발되면서 인공지능에서의 지식표현 (Knowledge Representation) 분야 연구의 가장 중요한 성과물로서 의료정보학 분야가 성공을 거두게 되었다 (Shortliffe, 1976). 이렇게 의료 분야가 인공지능 연구의 중요한 도메인이 된 이유는 오랜 세월을 거쳐오면서 의학지식은 체계화가 가장 잘 되었을 뿐만 아니라 지식을 규칙(rule)이라는 구조화되고 명시적인 형태로 표현하기가 용이하기 때문이다.

전문가시스템과 같은 지식베이스 시스템 (knowledge-based systems)에서의 지식은 대부분 다른 응용프로그램에서 사용되기 힘든 프로그램 종속적이며 한정된 상황에만 사용될 수 있도록 만들어진 것들이다. 이러한 한계를 극복하고 재사용 가능한 거대한 지식베이스를 개발할 필요성을 많은 연구자들이 인식하게 되었고, 새로운 형태의 지식표현 기술로서 온톨로지에 대한 연구가 본격적으로 시작되었다 (Lenat & Guha, 1990).

인공지능 분야와 상관없이 의료정보학에서 온톨로지 기술이 발전하게 된 계기가 된 것은 표준화된 용어시스템(terminology systems) 개발의 필요성에 대하여 공감을 갖게 되면서부터이다. 수많은 양의 환자 의무기록을 전자적으로 관리하고 협동적으로 환자 중심의 의료 서비스가 이루어지기 위하여 정보의 공유는 필수적이며, 이를 위하여 여러 의료기관 및 서비스 제공자들이 표준화된 용어체계를 사용할 필요가 있게 되었다. 또한 기존의 규칙 기반의 지식베이스 시스템들에서 표현하고 있는 지식 형태의 것들은 진정한 의미에서의 의학 지식은 아니라는 비판도 생기게 되었다 (Vander Lei & Musen, 1990). 의학 개념들과 개념들 사이의 관계들을 표준화된 용어 시스템 내에 표현할 수 있는 기술의 개발이 중요한 현안으로 부각되게 되었다.

위에서 언급하였듯이 의료정보학에서의 온톨로지 기술은 세 분야에서의 기술의 발전과정과 밀접한 연관을 맺고 있다. 즉 시맨틱 웹 구현을 위한 웹 응용기술, 인공지능의 지식표현 연구, 그리고 표준 용어시스템 개발이라는 서로 다른 전통을 통하여 자연스럽게 발전되었다. 본 논고에서는 의료정보학에서의 온톨로지 기술에 대한 중요한 이슈들을 이들 세가지 관점들을 동시에 고려하면서 살펴보도록 하겠다.

II. 온톨로지의 정의 및 특성

온톨로지의 특성을 가장 잘 표현하고 있는 Gruber(1993)가 내린 온톨로지에 대한 정의를 원문 그대로를 보면 다음과 같다. "An ontology is a formal, explicit specification of a shared conceptualization of a domain of interest." 앞에서 이미 언급한 내용이지만 다시 한번 요약해서 설명하자면, 여기서 형식적(formal)이라는 것은 기계가 읽고 처리할 수 있도록 술어논리와 같은 인공지능의 지식표현 방법으로 온

톨로지를 표현한다는 의미이고, 명시적(explicit)이라는 것은 개념들의 유형과 개념사용에 대한 규칙을 명시적으로 정의를 내려준다는 의미이다. 공유되었다(shared)라는 말은 용어의 사용이 관련 사람들의 합의에 의한 것이라는 뜻이고, 개념화(conceptualization)라는 말은 실제 세계에 대한 모형이라는 의미이며, 해당영역(domain of interest)이라는 말은 특정영역 중심이라는 뜻이다.

온톨로지의 원래 의미는 철학에서 존재론(Ontology, 대문자 'O'를 사용)을 뜻한다. 하지만 정보기술 분야에서의 온톨로지(ontology, 소문자 'o'를 사용하고 복수형인 ontologies도 가능)는 프로그램과 같은 일종의 설계된 인공물(designed artifact)이라 할 수 있다. 하지만 원래의 의미와 전혀 무관하지는 않다. 즉, 논리적 관계로만 표현된 지식이 존재에 대하여 아무 얘기도 못하는 것과는 달리 온톨로지는 사물이나 사건이 존재하는 세계의 구조를 표현하고자 하는 것이다.

온톨로지를 구축하는 것은 해당영역의 사람들의 동의만 있으면 가능한데 이를 위하여 협력적 과정이 필요하다. 온톨로지는 일반화의 정도에 따라서 다음과 같이 몇 가지 유형별로 나뉘어 질 수 있다.

- 도메인 온톨로지(domain ontology): 특정 분야에 한정되는 온톨로지(예: 의료, 전자, 기계 등)
- 과업 온톨로지(task ontology): 특정 과업을 수행하기 위하여 개념들을 정의한 온톨로지(예: 진단, 판매)
- 메타데이터 온톨로지(metadata ontology): Dublin Core와 같이 온라인 정보자원의 내용을 기술하기 위한 어휘를 제공
- 상식개념 온톨로지(common sense ontology): 세계에 대한 상식적 개념을 표현 (예: 시간, 공간, 움직임 등)
- 표현을 위한 온톨로지(representational ontology): 어떤 특정영역에서의 개념에 관한 것이라기보다는 다른 개념을 온톨로지로 표현하는데 쓰이는 개념에 관한 것 (예: Frame Ontology는 여러 개념을 객체지향적으로 표현할 때 쓰이는 프레임, 슬롯 등과 같은 객체지향적 개념을 표현하는 메타수준의 온톨로지)

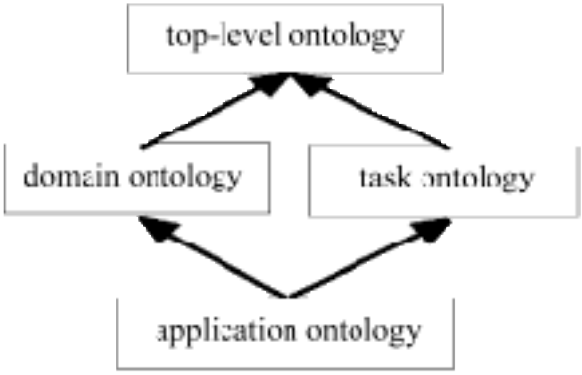

Fig 1. 온톨로지들 사이의 관계 출처: Sowa, 2000

III. 지식표현과 온톨로지

인공지능의 지식표현(knowledge representation) 기술의 구현을 위한 대표적 접근방법으로 논리적 방법과 구조적 방법이 있다. 논리적 방법 중에서 그 동안 가장 많은 사람들의 관심의 대상이 되었던 것이 일차술어논리(first-order predicate logic)이다. 일차술어논리는 비교적 적은 수의 원초적 기호들(primitives)만을 포함하고 있으면서도 표현력(expressiveness)이 매우 풍부하고, 잘 정의된 모형 이론(model theory) 및 증명 이론(proof theory)을 갖고 있다. 논리적 지식표현 방식은 잘 정의된 추론의 절차를 기계가 처리 가능한 언어로 표현할 수 있을 뿐만 아니라 모순과 모호함이 없는 수학적 엄격성을 보장해준다는 장점이 있다. 따라서 대부분의 규칙 기반 전문가 시스템들은 일차술어논리를 이용한 논리 프로그래밍 언어로 구현되었다.

논리 기반의 지식표현 언어들을 사용하면, 지식은 상대적으로 규모가 작은, 독립적인 단위 요소들(chunks)로 조직화된다. 직관적으로 볼 때, 이러한 방식이 지식을 조직화하는 가장 자연스러운 그것은 아니다. 왜냐하면 적어도 사람을 생각해 볼 때, 사람 마음 속의 지식이란 사실들(facts)의 단순한 집합에 불과한 것은 아닌 것 같기 때문이다. 오히려 지식은 개념적인 실체들(conceptual entities)의 집합으로서, 그것들과 연관된 기술들(associated descriptions)이 부가된 상태로, 보다 규모가 큰 단위 요소들을 써서 저장되는 것처럼 보인다. 게다가 이 경우에는 지식의

다양한 단위 요소들 사이에서 대규모의 연결을 상정하게 되는 것이 보통이다. 구조적(structural) 혹은 인지적(cognitive) 접근 방식은 지식 표현에 있어 이상과 같은 우리의 직관을 반영한다. 의미망(semantic network)과 프레임(frame) 기반의 언어는 이것의 대표적인 예이다. 둘 사이의 차이는 의미망에는 상호연결성(interconnectedness)이라는 직관이 보다 직접적으로 투사된 데 비하여, 개념적인 실체들에 연관된 기술(description)들이란 직관은 프레임 기반의 언어에서 보다 명확하게 시각화된다는 점에 있다. 최근 온톨로지 표현 언어로 가장 주목을 받고 있는 기술논리(Description Logics)는 지식표현의 논리적 접근방식의 엄격성과 구조적 접근방식의 직관성을 통합시킨 방법이라 할 수 있다. W3C에서 발표한 표준 웹 온톨로지 표현 언어인 OWL도 기술논리에 근간을 두고 있다.

온톨로지도 넓은 의미에서는 데이터베이스라고 할 수 있지만 데이터보다 복잡한 형태의 지식과 관련되어 있다는 의미에서는 지식베이스(Knowledge-base)라고 부르기도 한다. 하지만 엄격히 말해서 온톨로지는 지식 내용과 절차적 추론과정을 포함하는 포괄적 의미의 지식보다는 용어들 사이의 '개념'적 관계에 국한되기 때문에 지식베이스와 구별되는 또 다른 형태의 데이터베이스라고 할 수 있다. 지식베이스가 특수한 개별적 사실에 대한 표현을 위한 것이라면 온톨로지는 보편적 개념을 표현하는 것으로 여러 목적에 재사용 가능한 것이라 할 수 있다.

지식의 표현은 어휘의 사용을 통하여 가능하다. 어휘의 사용은 두 가지 형태로 구분될 수 있다. 하나는 어휘의 내용을 다른 어휘를 사용하여 명확하게 하는 것이고, 다른 하나는 어휘들로 표현된 사실들 사이의 논리적 관계를 통하여 새로운 사실을 알아내는 것이다. 전자를 개념(concept)에 대한 표현이라고 하고 후자를 추론(inference)이라고 한다. 논리적 기법을 사용한 기존의 지식베이스시스템에 대한 연구는 표현력이 풍부하면서도 계산력도 우수한 지식표현 방법을 개발하는 것으로서, 추론이라는 절차적 의미의 지식표현을 개념의 표현보다는 더 비중을 두거나 아니면 모두 같은 방식으로 취급한다. 따라서 어휘의 개념 지식들을 포함하고 그것들이 적용되는 영역(domain)에 대한 고려는 하지 않고 있다. 하지만 온톨로지는 절차적 지식에 대한 표현보다는 어떤 영역에서 공통적으로 사용되고 있는 어휘들의 집합을 개념적으로 표현하는 방법이라고 할 수 있다.

예를 들어 심장질환을 진단하는 전문가시스템을 만든다 라고 가정하자. 전문가시스템은 환자의 건강과 관련된 데이터를 사용하여 논리적 추론과정을 거쳐 병을 진단하는 의사의 지식을 지식표현을 위한 컴퓨터언어로 구현한 프로그램이다. 전문가시스템에서는 추론과정을 논리적으로 표현하고 몇몇 전문 용어들 사이의 관계를 논리적으로 표현하고 있지만, 심장병 전문의들이면 누구나 공통적으로 사용하고 있는 용어들에 대한 정의를 표현하고 있지는 않다. 이에 반하여 온톨로지는 어휘의 정의를 다른 어휘와의 논리적 관계뿐만 아니라 '혈액', '혈압', '심장' 등과 같은 가장 기본적(primitive) 어휘에서부터 지어져 나가는(bottom-out) 구조를 통하여 나타낸다. 이는 지식표현이 해당분야의 전문가들이 동의하거나 공통적으로 사용하는 가장 기본적인 어휘에서부터 출발한다는 점에서 해당분야 특정적(domain specific)이라 할 수 있다. 온톨로지는 해당분야의 전문가들의 관심, 목적, 그리고 관점을 반영하며, 따라서 온톨로지의 역할은 해당분야에서 사용되는 어휘의 해석과 의미의 한계를 정하는데 있다. 이러한 의미에서 온톨로지는 사용되는 어휘들간의 논리적 형식보다는 어휘들의 내용을 중심으로 창조 된다고 할 수 있다.

온톨로지는 기존의 관계형 데이터베이스 이론의 개념 스키마와도 구분되는바, 개념 스키마의 경우에 있어서는 각각의 개념 스키마, 관계형 데이터베이스, 그리고 그것들의 속성들의 의미가 무엇인지에 대하여 아무런 언급을 하지 않고도 단순히 패턴매칭의 방법으로 필요한 정보를 이끌어 낸다. 온톨로지는 해당분야의 전문가들 - 인간 혹은 컴퓨터 에이전트들 - 사이에서 그 의미를 동의하는 가장 기본적인 어휘에서부터 출발하기 때문에 전문가들의 협동작업이나 분산처리 시스템에서의 정보교환 등을 위하여 새롭게 내용 수준의 정의를 반복할 필요가 없다. 즉, 온톨로지들 사이에 용어의 사용이 다르고 개념들의 관계적 논리 구조가 다르다고 할지라도 정보의 처리방식이 단순한 패턴매칭이 아닌 내용 중심적이기 때문에 정보의 통합과 공유가 가능하다. 따라서 이러한 온톨로지에 대한 연구는 지식의 공유(share)

와 재사용(reuse)을 가능하게 하는 웹과 같은 거대한 지식기반시스템 개발과 깊은 연관을 맺을 수 밖에 없다.

Ⅳ. 용어시스템과 온톨로지

컴퓨터 기반의 환자 기록(computer-based patient records, CPRs)의 형태로 의료 데이터를 저장하고 사용하기 시작하면서 구조화되고 통제된 데이터 표현의 필요성이 증가하게 되었다 (Moorman et. al., 1995). 따라서 표준화된 용어들(terms)의 체계적 사용을 가능하게 하는 용어 시스템(terminology systems)은 의료정보학의 중심 주제가 되었다. 의료 분야의 용어 시스템들은 사용의 목적, 기능, 구조의 차이에 따라서 "classification", "thesaurus", "vocabulary", "nomenclature", "ontology" 등의 이름으로 불려지고 있다 (Keizer et. al., 2000).

일종의 언어학적인 표시(label)인 '용어(term)'는 사물에 대한 '개념(concept)'을 지칭할 때 사용되고, '개념'과의 매핑 관계를 통하여 그 의미가 획득된다. 의료 용어 시스템에서 '코드(code)'란 컴퓨터 관점에서의 '용어'라고 할 수 있으며, 주로 문자, 숫자, 혹은 이 둘의 결합으로 이루어진다. '코드'는 크게 의미를 가지는 코드인 'significant code'와 의미를 가지지 않는 'non-significant code'로 구분된다. 전자는 용어 시스템에서 개념을 지칭하기 위한 것으로 주로 사용자가 이해하기 쉬운 형태의 'mnemonic codes' (예: A900 = Anatomical component "lung")나 개념의 구조를 표시하기 쉬운 'hierarchical codes' (예: 480 = viral pneumonia 이고 480.1 = adeno virus) 등으로 되어있다. 후자는 용어의 개념과는 무관하게 용어의 유일성을 보장하기 위한 'random codes'나 'sequential codes' 형태로 되어있다.

용어 시스템에서 개념을 표현하기 위하여 기본적으로 각 개념의 속성과 개념들간의 관계들을 나타낼 수 있어야 한다. 예를 들어, "위 궤양" 개념의 속성을 나타내기 위하여 궤양의 '색', '크기', '모양' 등이 기술되어야 한다. 개념들간의 관계는 위계적(hierarchical) 관계와 비위계적(nonhierarchical) 관계로 표현될 수 있다. 즉, 위계적 관계는 부분(partitive) 관계나 일반화(generic) 관계같이 개념들간의 순서를 표현하는 것이고, 비위계적 관계에는 개념들 사이에 위계적 구조를 가질 수 없는 공간적, 시간적, 인과적 관계 등이 포함된다. 개념을 좀 더 풍부하게 표현하고 형식화하는 온톨로지적인 표현 방식을 의료 용어 시스템에 도입하려는 시도가 활발히 이루어지고 있다.

최근에 개발된 대표적 의료 용어 시스템들로 ICD-9-CM/ICD-10, SNOMED, NHS Clinical Terms, UMLS, GALEN 등이 있다. 온톨로지적인 관점에서 ICD의 가장 큰 문제점은 개념의 정의와 개념들간의 관계가 명시적(explicit)으로 표현되고 있지 않다는 점이다. 또한 개념과 용어를 분리시키고 있지 않으므로 동의어의 관계가 표현될 수 없으며, 기존의 개념들로부터 새로운 개념을 합성할 수도 없다. NHS Clinical Terms, UMLS, SNOMED는 이러한 문제점들은 가지고 있지 않으나, 개념을 정의하거나 합성 개념을 만드는 형식적(formal) 규칙은 없다. UMLS는 의미망 기반으로 구성되어 개념을 구조화 하고 있으나 추론 기능은 가지고 있지 않다.

GALEN은 다른 용어 시스템들과는 다르게 가장 온톨로지적인 개념 구조를 표현하고 있는 시스템이다. 즉, GALEN의 온톨로지 표현 언어인 GRAIL은 기술논리(Description Logics) 기반의 형식 규칙을 제공함으로써 풍부한 개념 표현의 방법을 제공하고 있다. 온톨로지 형태의 용어 시스템의 가장 중요한 기능은 재사용 가능한 개념으로 의학 지식을 표현하고 있다는 점이다. GALEN의 CORE Model은 해부학, 생리학, 병리학에 있어서의 가장 핵심이 되는 의학 개념들을 포함하고 있으며 이를 특수한 상황에 재사용할 수 있는 방법을 제공하고 있다. GALEN에서 개념의 재사용을 위한 구조를 Rector(2002)는 "느슨하게 연결된(loosely coulpled) 개발 환경"이라고 부른다.

Ⅴ. 의료 온톨로지 기술의 현안

1. 의료 온톨로지의 규모

온톨로지의 구축에 있어서 중요한 조건이 되는 것이 표현력(expressivity)과 계산의 용이성(tractability)이라는 이율배반적 특성을 동시에 고려해야만 한다. 어떻게 서로 모순되는 두 가지 조건을 충족시키기

위하여 고려해야 할 요소가 바로 온톨로지의 규모(scale)라 할 수 있다. Rector는 규모의 문제를 다음의 세부 현안들로 나누어 설명하고 있다 (Rector, 2002).

- 크기(size): 의료 온톨로지에서 현실적으로 사용 가능한 최소한의 개념의 수는 10,000개이며, 종합적으로는 250,000 정도라 할 수 있다. 개념의 수가 많으면 표현력에서는 우수하나 추론 절차는 그만큼 복잡하게 된다.
- 연결성(connectivity): 의료 온톨로지에서의 개념들은 서로 연결성이 우수해야 한다. 대부분의 해부학적인 개념들은 부분·전체 관계로 서로 연결되어 있고, 서로 다른 분야의 의학 개념들도 기능적, 인과적 관계들로 연결되어 있다. 연결성이 높으면 표현력이 우수하지만 반대로 계산은 용이하지가 않다.
- 정밀도(granularity): 온톨로지가 구축되는 분야에 따라서 개념 표현의 정밀도가 달라질 수 있다. 예를 들어, 미생물학, 생리학, 심리학에서 사용되어지는 온톨로지의 정밀도는 다르다.
- 복잡성(complexity)과 상세도(levels of details): 온톨로지는 용어 체계와 달리 개념을 복잡하고 상세하게 표현할 수 있다. 하지만 어느 정도 상세하게 표현되어야 하는 문제는 온톨로지의 사용 목적에 따라서 달라질 수 있다.

2. 의료 온톨로지의 표현력

의료 온톨로지는 개념의 일반화 관계(IS-A 혹은 Kind-Of) 외에 부분 관계(partitive or mereological relation)의 표현이 매우 중요하다. 일반적으로 개념들 사이의 부분 관계도 일반화 관계와 마찬가지로 이행적 특성(transitivity)을 가지고 있지만, 속성의 상속(inheritance) 문제에 있어서 항상 그런 것은 아니다. 부분관계에서의 상속의 방향성은 일반화 관계와는 달리 경우에 따라서 양 방향이 다르게 나타난다.

- 하향전이(downward distribution): 전체 개념의 속성이 부분 개념의 속성으로 전이되는 현상을 뜻한다. 예를 들어, 자동차가 주차장에 위치해 있으면, 자동차의 부분인 엔진도 주차장에 위치한다.
- 상향전이(upward distribution): 부분 개념의 속성이 전체 개념의 속성으로 전이되는 현상을 뜻한다. 예를 들어, "랑게르한스 섬"의 이상은 췌장에 이상이 있는 것으로 인식될 수 있다.

질병과 해부학적 기관과의 관계에 있어서 상향전이가 적용되는 것이 보통이지만 항상 그런 것은 아니다. 예를 들어, 랑게르한스 섬은 췌장의 부분이고 췌장은 소화기계통의 부분이지만 상향전이는 두 단계로 올라가지는 않는다.

부분관계의 복잡성으로 인해 온톨로지 개념들간의 부분관계의 표현은 매우 어려운 문제이다. GALEN 프로젝트에서는 Winston(1987)과 Odell(1998) 등이 제안한 아래의 부분관계들을 기본으로 하여 여러 종류의 부분관계를 표현하고 있다.

- Component: 예로서 '자동차 문 손잡이'
- Stuff: 예로서 '빵의 재료인 밀가루'
- Portion: 예로서 '빵의 한 조각'
- Area: 예로서 '시에 속해 있는 구'
- Member: 예로서 '함대에 속해 있는 군함 한 척'
- Partner: 예로서 '로미오와 줄리엣에서의 로미오'
- Piece: 예로서 '문에서 분리된 손잡이'

부분관계의 표현 외에도 기능관계, 인과관계, 변화를 나타내는 시간관계 등의 표현도 의료 온톨로지의 구축에 있어서 고려되어야 할 중요한 문제들이다.

3. 의료 온톨로지의 응용

의료 온톨로지가 사용되는 영역은 다양하다. 우선, 온톨로지를 통한 의미적 개념 표현은 정보의 검색을 도와준다. 즉, 환자의 관점에서 의료 서비스 제공자는 환자에게 가장 적절한 정보를 찾아줄 수 있다. 정보의 검색과 같은 맥락에서 생각해 볼 수 있는 응용 영역은 진단과 같은 의사결정(decision making)과 자동 인덱싱에 응용될 수 있다. 기존의 지식베이스와는 다른 차원에서 풍부한 개념들간의 관계 표현을 통하여 보다 우수한 의미적 정보처리가 가능할 수 있게 되었다.

온톨로지가 문제상황에 적용될 수 있기 위해서는 우선 다양한 규모와 내용의 온톨로지들을 저장하고 있는 라이브러리에서 가장 적절한 온톨로지를 발견

하고 필요에 따라서 여러 온톨로지들을 결합하고, 변형시키고, 통합할 수 있어야 한다. 이를 위하여 온톨로지 통합을 위한 여러 자동화 및 반자동화 된 툴의 개발이 중요하다(Noy, 2002).

온톨로지가 재사용 가능한 형태로 되어 있듯이 최근에 도메인 온톨로지를 사용하여 주어진 문제상황에 가장 적절한 해결방안을 찾는 방법에 대한 연구도 이루어지고 있다. EON 프로젝트는 소위 재사용 가능한 다양한 "problem solving methods"들의 라이브러리를 구축하고 있다. 그 한 예로 EON의 'therapy-determination problem-solving method'의 "plan" 인자는 환자에게 적합한 치료방법을 발견해주기 위하여 도메인 온톨로지의 "guideline" 개념에 매핑이 된다(Musen, 1998).

VI. 결 론

의료 분야의 지식의 재사용성을 높이고, 환자의 의무 기록과 같은 엄청난 양의 정보에 대한 자동화된 처리, 그리고 궁극적으로 환자 중심의 의료 서비스 제공을 위하여 온톨로지 기술은 필수적이다. 본고에서는 의료정보학에서의 온톨로지 기술의 토대가 된 시맨틱 웹 구현을 위한 웹 응용기술, 인공지능의 지식표현 연구, 그리고 표준 용어 시스템 개발이라는 세 가지 기술적 관점에서 살펴봄으로써, 이 분야 연구의 중요한 현안들에 대한 이해를 돕고 지금까지의 의료정보학의 연구 성과들을 온톨로지 기술과 연결할 수 있는 가능성에 대하여 고찰하였다.

참고문헌

1. Gruber, T.R. A Translation Approach to Portable Ontology Specifications, Knowledge Acquisition, 1993; 5: 199-220
2. Keizer, N.F., Abu-Hanna, A., & Zwetsloot-Schonk, J.M.M. Understanding Terminological Systems I: Terminology and Typology. Methods of Information in Medicine, 2000; 39: 16-21
3. Lenat, D.B. & Guha, R.V. Building Large Knowledge-Based Systems, Addison Wesley, Reading, MA, 1990
4. Ledley R.S & Lusted L.B. Reasoning foundations of medical diagnosis. Science, 1959; 130: 9-21
5. Musen, M.A. Domain Ontologies in Software Engineering: Use of Protege with the EON Architecture. Methods of Information in Medicine, 1995; 34: 85-95
6. Odell, J.J. Advanced Object-Oriented Analysis and Design Using UML. Cambridge University Press, 1998
7. Rector, A. Medical Informatics, in Baader et. al. (eds) The Description Logic Handbook: Theory, Implementation, and Applications, 2003
8. Shortliffe, E.H. Computer-Based Medical Consultations: MYCIN, American Elsevier, New York, 1976
9. Sowa, J.F Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks/Cole Press, 2000
10. Winston, M.E., Chaffin, R., & Herrmann, D. A Taxonomy of Part-Whole Relations. Cognitive Science, 1987; 11: 417-444