

傷寒論 온톨로지 구축 방법론 연구

정태영 · 김희열¹ · 박종현^{1*}

대구한의대학교 한의과대학 진단학교실, 1: 병리학교실

Study on a Methodology for Developing Shanghanlun Ontology

Tae Young Jung, Hee Yeol Kim¹, Jong Hyun Park^{1*}

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

Knowledge which is represented by formal logic are widely used in many domains such like artificial intelligence, information retrieval, e-commerce and so on. And for medical field, medical documentary records retrieval, information systems in hospitals, medical data sharing, remote treatment and expert systems need knowledge representation technology. To retrieve information intellectually and provide advanced information services, systematically controlled mechanism is needed to represent and share knowledge. Importantly, medical expert's knowledge should be represented in a form that is understandable to computers and also to humans to be applied to the medical information system supporting decision making. And it should have a suitable and efficient structure for its own purposes including reasoning, extendability of knowledge, management of data, accuracy of expressions, diversity, and so on. we call it ontology which can be processed with machines. We can use the ontology to represent traditional medicine knowledge in structured and systematic way with visualization, then also it can also be used education materials. Hence, the authors developed an Shanghanlun ontology by way of showing an example, so that we suggested a methodology for ontology development and also a model to structure the traditional medical knowledge. And this result can be used for student to learn Shanghanlun by graphical representation of it's knowledge. We analyzed the text of Shanghanlun to construct relational database including it's original text, symptoms and herb formulars. And then we classified the terms following some criterion, confirmed the structure of the ontology to describe semantic relations between the terms, especially we developed the ontology considering visual representation. The ontology developed in this study provides database showing fomulas, herbs, symptoms, the name of diseases and the text written in Shanghanlun. It's easy to retrieve contents by their semantic relations so that it is convenient to search knowledge of Shanghanlun and to learn it. It can display the related concepts by searching terms and provides expanded information with a simple click. It has some limitations such as standardization problems, short coverage of pattern(證), and error in chinese characters input. But we believe this research can be used for basic foundation to make traditional medicine more structural and systematic, to develop application softwares, and also to applied it in Shanghanlun educations.

Key words : shanghanlun, ontology, protege

서 론

컴퓨터가 우리 사회에 등장한지 불과 수십 년밖에 되지 않았지만 IT산업의 급속한 발달로 현대사회를 흔히 '정보시대' 또는 '정보사회'라 일컬을 만큼 컴퓨터와 정보통신망을 이용한 정보 창출과 이를 이용한 가치생산이 사회를 움직이는 원동력이

되고 있다¹⁾. 지식정보를 지능적으로 검색하고 질적으로 향상된 정보서비스를 위해서는 지식을 기술하고 표현하며 공유할 수 있도록 하는 체계적이며 통제된 메커니즘이 필요하다²⁾. 특히, 의사결정을 지원하는 전문가시스템 등 의료영역에 고급적으로 활용하기 위해서는 의학 전문지식을 컴퓨터와 사람이 동시에 이해할 수 있는 형태로 나타내어야 하며, 목적달성에 부합되는 구조

* 교신저자 : 박종현, 대구시 수성구 상동 165, 대구한의대학교 한의과대학

· E-mail : moguri@dhu.ac.kr, · Tel : 053-770-2248

· 접수 : 2011/08/31 · 수정 : 2011/09/20 · 채택 : 2011/09/26

1) 김선일, 김덕원, 박광석, 차은중, 김남균, 이상훈. 의료정보학. 여문각. 2008:10.

2) 노상규, 박진수. 인터넷 진화의 열쇠 온톨로지. 가즈토이. 2007:9.

를 가져야 하고 추론의 효율성, 지식 획득의 용이성, 저장의 간결성 및 표현의 정확성, 다양성 등을 갖추어야 하는데³⁾, 이와 같은 기계가 해석 가능한 지식 체계를 온톨로지(Ontology)⁴⁾라고 부른다⁵⁾).

최근 한의학 분야에서도 온톨로지 기술을 도입한 지식표현과 한의 지식의 효과적인 활용방안에 대한 다양한 연구가 진행되어, 이기는 온톨로지의 개념모델에 의한 한의학 처방지식관리시스템을 구현하였으며, 문 등⁶⁾과 박 등⁷⁾은 진단을 위한 온톨로지 연구 결과를 보고하였다. 특히 한의학연구원의 의료정보 연구 그룹에서는 의료정보기술을 활용한 각종 연구¹⁰⁾¹¹⁾¹²⁾¹³⁾와 사업들을 추진하고 있다. 온톨로지 기술을 이용하여 한의 지식을 구조화, 체계화하여 시각적으로 표현한다면, 화면상에서의 연계 검색은 물론 학습 자료로도 활용 가능성이 크다.

이에 저자는 상한론에 기재된 지식의 온톨로지를 시범적으로 구축함으로써 한의학의 온톨로지 구축방법론을 제시함은 물론 한의 지식 구조화에 하나의 모델을 제시하고, 더 나아가 그래프를 통하여 상한지식을 시각적으로 표현함으로써 상한론 학습에 도움을 주고자 본 연구를 시도하였다.

본 연구에서 저자는 상한론에 기재된 조문원문, 처방, 탕증, 증상 및 처방 구성약물 상호 간의 관계형 데이터베이스 구축을 위하여 온톨로지 구조를 결정하고 각 용어의 속성을 규정하여 이들 상호간의 관계를 설정하였으며, 특히 시각적 표현에 염두를 두고 온톨로지를 구축하였다.

본 론

1. 연구방법

1) 온톨로지 구축을 위한 도구 선정

(1) 온톨로지 표현 언어

본 연구에서는 표현력이나 추론 능력에 있어서 뛰어난 시맨틱 웹 온톨로지 언어로 평가되고 있으며 또한 널리 사용되고 있는 OWL(Ontology Web Language)을 사용하였다. OWL은

DAML+OIL에 기반을 둔 온톨로지 구축 경험을 토대로 개념의 일관성을 확보하여 클래스와 속성의 개념 및 그들 사이의 관계가 보다 명료하게 정의되도록 정리한 온톨로지 언어로 W3C에서 표준화하고, 2004년 2월에 발표된 온톨로지 마크업 언어이다. OWL은 완전히 새로운 언어로서의 의미보다는, XML, RDF, RDFS와 같은 기존의 마크업 표준으로부터 시맨틱 웹으로의 진행을 위해 필요한 사항을 추가적으로 정의한 것이며¹⁴⁾, DAML+OIL처럼 클래스와 하위 클래스, 속성과 하위 속성, 속성 제한, 클래스와 속성 개체를 가지고 있으며, 대부분의 DAML+OIL 어휘를 이해하기 쉽게 변경하였다¹⁵⁾.

(2) 온톨로지 개발 도구

온톨로지를 쉽게 개발하고, 관리하고 공유하며, 여러 가지 서비스에 응용하도록 지원하는 툴은 매우 중요하다고 볼 수 있다. 새로운 온톨로지를 구축할 수 있는 온톨로지 개발 툴에는 Ontolingua 서버, OntoSaurus, WebOnto, OilEd, SemanticWorks™ 2006 등 지식모델이 특정 온톨로지 언어에 종속되어 있는 툴과, Protege, WebODE, OntoEdit 및 KAON 등 언어에 독립적인 툴이 있으며, 본 연구에서는 Protege¹⁶⁾를 사용하였다.

2) 상한론 판본(版本) 선정과 용어의 주요 근거 서적 선정

상한론은 편찬 직후 散失되어 원본이 전해지지 않고, 다만 여러 서적에 수록되어 傳本도 다양하여 원서의 정확한 체재와 면모를 확인할 수는 없다. 현재 상한론 의가들에게 인정되고 있는 상한론 판본은 宋本, 康平本과 康治本 3종이 있는데, 본 연구에서는 3종의 판본 중 가장 많이 연구되고, 전국 한의과대학의 주 교재로 사용되고 있는 송본상한론을 중심으로 온톨로지를 구축하였다. 또한 같은 송본 상한론이라고 하더라도 글자나 방점 등 약간의 出入이 있는데, 본 연구에서는 한의과대학에서 상한론 교재로 사용하고 있는 『현대상한론』¹⁷⁾에 기재된 상한론 조문을 중심으로 온톨로지를 구축하였다.

2. 연구 진행 과정

1) 온톨로지 구조 결정

(1) 온톨로지 영역 설정

본 연구에서 구축하고자 하는 온톨로지는 일차적으로 상한론 조문에 나타난 내용을 구조화하여 데이터베이스화 하는 것이었다. 상한론 조문은 ‘太陽病, 頭痛發熱, 汗出惡風, 桂枝湯主之.’¹⁸⁾처럼 기본적으로 환자의 상태를 언급하고 主治 처방을 제시한 구조로 되어있으며, 처방 내용은 조문 뒤에 따로 서술하는 방식을 취하고 있다. 따라서 본 연구의 온톨로지도 ‘증상’, ‘처방’, ‘구성약재’를 중심으로 이들의 관계를 규정짓는데 주안점을 두고 Fig. 1에서 보는 것처럼 이들 간의 관계를 표현할 수 있도록 설계

3) 이재규, 최형림, 서민수, 주석진, 지원철. 전문가시스템 원리와 개발. 법영사. 2002:28-9.

4) 온톨로지는 철학의 존재학에서 유래된 것으로(두산백과사전 EnCyber & EnCyber.com), 구성원의 동의에 의해 수용되는 개념과 개념들 간의 관계를 표현한 지식을 말한다.

5) 임희숙. 온톨로지 기반 전자진료부 입력 인터페이스 설계 및 구현. 경희대학교 대학원 동서의료공학과 석사학위논문. 2004:5.

6) 박경모, 임희숙, 박종현. Protege를 이용한 한의학의 구조화된 증상입력을 위한 온톨로지 개발. 동의생리병리학회지. 2003;17(5):1151-6

7) 이현실. 합성 온톨로지 기반의 한의학 처방 지식관리 시스템. 한국학술정보(주). 2006.

8) 문경실, 박수현. 맞춤형 진단 서비스를 위한 한의학 온톨로지. 한국컴퓨터정보학회논문집. 2010;15(1):23-30.

9) 신상우, 정길산, 박경모, 김선호, 박종현. 한의진단 Ontology 구축과 평가. 동의생리병리학회지. 2006;20(1):202-8.

10) 김 철, 김상균, 송미영. 논문분석과 구축사례 조사를 통한 한의학 온톨로지 연구동향 분석. 한국한의학연구원논문집. 2008;14(2):121-9.

11) 김상균, 김 철, 장현철, 예상준, 송미영. 약재 온톨로지를 활용한 약재 검색 매쉬업 시스템. 정보관리연구. 2008;39(4):173-86.

12) 송미영. 전통약재 정보의 효율적인 활용을 위한 온톨로지 구축. 한국한의학연구원 보고서. 2007.

13) 송미영. 온톨로지 기반 한의학 지능형 정보체계 연구(1차년도). 한국한의학연구원 보고서. 2009.

14) Deborah L. McGuinness. Frank van Harmelen. 2004. OWL Web Ontology Language Overview. W3C.

(<http://www.w3.org/TR/owl-features/>)

15) 노상규 등. 앞의 책 : 116.

16) 현재 v.4.1까지 개발되었으나 본 연구에서는 protege 3.31을 사용하였다.

17) 문준전 등. 현대상한론. 한의문화사. 2005.

18) 송본상한론 '변태양병맥증병치 상' 12번 조문.

하였다.

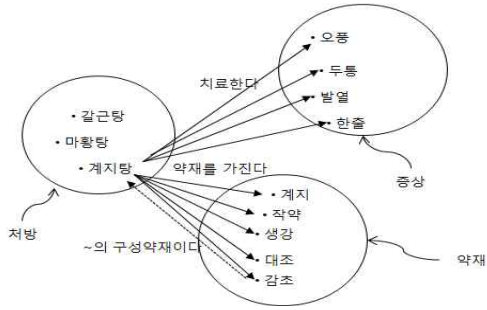


Fig. 1. 상한론 온톨로지의 주요 영역

(2) 매개 개념의 도입

본 연구에서 추구하는 목표 중의 하나는 상한론 지식을 그래프로 표현하여 학습에 도움을 주고자 하는 것이다. 만약 계지탕과 관련된 모든 정보를 계지탕과 직접 연결하여 그래프로 표현한다면 관계 그래프가 지나치게 복잡할 뿐만 아니라 내용을 체계적으로 파악하기가 쉽지 않을 것이다. 따라서 본 논문에서는 이러한 번잡함을 피하기 위하여 매개개념을 도입하여 간략하게 묶어 정리하는 형식을 취하였다. 위 Fig. 1에서 '계지', '작약', '생강', '대조', '감초' 등 계지탕을 구성하는 약재에 대하여 각각 약재를 계지탕과 직접 연결하지 않고 '계지탕 구성약물'이라는 매개개념을 사용하여 표현하였으며, 계지탕이 치료하는 각 증상도 '계지탕증'이라는 항목을 설정하여 계지탕은 '계지탕 구성약물'이라는 처방전을 가지며 '계지탕증'을 치료하는 것으로 표현하였다(Fig. 2). Fig. 2는 Fig. 1에 비하여 계지탕에 직접 연결된 고리가 훨씬 적어 간결한 장점이 있다.

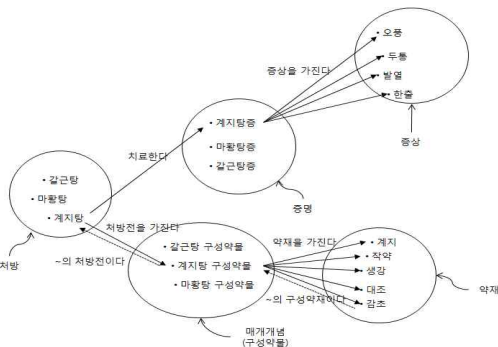


Fig. 2. 매개개념의 도입

(3) 온톨로지 구조

이상에서 말한 '처방', '증상' 및 '약재'가 상한론 내용을 이루는 주요 골자이지만, 이들을 효과적으로 표현하기 위하여 Fig. 3과 같은 구조의 온톨로지를 설계하였으며, 조문 원문도 함께 제공하고 관련 처방과의 연결을 모색하였다.

2) 개념 간 관계 설정을 위한 속성(Properties) 규정

속성은 개념이 근본적으로 가지고 있는 성질을 말하는데, 클래스나 인스턴스의 특정한 성질, 성향 등을 나타내기 위하여 클래스나 인스턴스를 특정한 값(value)과 연결시킨 것이다.

모두 '이름'을 가지고 있으며, '색깔', '맛', '생산지' 등과 같은 속성을 가지고 있다. 예를 들어 "홍화의 색은 빨간색이다"라는 것을 표현하기 위하여 'hasColor'와 같은 속성을 정의할 수 있으며, 약재의 이름을 표현하기 위하여 'hasName'과 같은 속성을 정의할 수 있을 것이다. 상한론 온톨로지 구축을 위하여 본 연구에서는 앞의 Fig. 3과 같은 온톨로지 구조를 결정하고 각 개념의 관계 설정을 위하여 21개의 속성을 설정하였다¹⁹⁾.

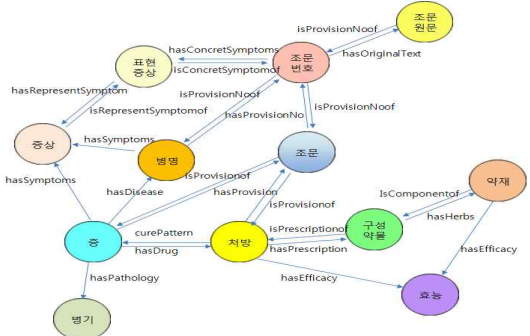


Fig. 3. 상한론 온톨로지 구조

3) 처방, 구성약물, 증상 용어 추출과 정리

(1) 처방의 정리

처방의 정리는 『傷寒來蘇集』²⁰⁾의 그것처럼 비슷한 부류의 처방을 함께 묶어 정리하는 방식을 취하였다. 본 연구에서는 상위 부류의 수를 적게 하기 위하여 『傷寒論手冊』²¹⁾의 분류를 따라 『傷寒論』에 수록된 113개의 처방을 桂枝湯類(19방), 麻黃湯類(6방), 葛根湯類(3방), 瀉心湯類(11방), 柴胡湯類(6방), 梔子豉湯類(7방), 白虎湯類(3방), 承氣湯類(12방), 五苓散類(4방), 四逆湯類(11방), 理中湯類(9방), 기타(22방)로 분류하였다.

(2) 구성약물 정리

상한론에 수록된 처방을 구성하고 있는 약재는 현재 전국한의과대학에서 공통교재로 사용하고 있는 본초학 교재²²⁾를 중심으로 분류하였으며, 본초학 교재에 기재되어 있지 않은 약재는 효능을 중심으로 임의 분류하였다.

(3) 증상 용어 추출과 정리

증상 용어는 『現代傷寒論』²³⁾의 색인에서 추출하였다. 복합증상은 분리하여 각각의 증상으로 처리하였으며, 유사증상의 처리를 위하여 비슷한 의미의 증상을 한데 묶어 대표증상을 정하고 나머지는 유사증상으로 처리하여 정리하였고, 온톨로지 구축시 대표증상과의 연결은 물론 각 탕증의 관련 조문과 관계를 설정하여 검색이 가능하도록 하였다. 증상은 『中醫症狀鑑別診斷學』을 따라 분류하였다.

4) 용어 입력과 관계 설정

19) 본 연구에서 정한 속성의 이름은 임의로 정한 것이며, 앞으로 논의할 온톨로지 연구자들의 합의하에 통일된 명칭을 정해야 할 것이다.

20) 柯琴. 傷寒來蘇集. 중국중의학출판사. 1998.

21) 張啓基, 王輝武, 郭自光. 傷寒論手冊. 과학기술문헌출판사중경분사. 1983.

22) 전국한의과대학 본초학공동교재 편찬위원회. 본초학. 영림사. 2010.

23) 문준전 등. 현대상한론. 한의문화사. 2005:671-9.

이상에서 설정한 온톨로지의 구조, 속성과 정리된 용어는 본 연구에서 온톨로지 구축도구로 선택한 Protege 3.3.1을 이용하여 입력하고 관계를 설정하였다. 각 용어의 漢字 입력에 관한 문제는 일부 漢字에서 오류가 발생하여 특별한 경우를 제외하고는 우선 한글로 입력하였다. 우선 Protege의 'Classes' 탭에서 본 연구에서 다루는 영역과 관련된 개념과 분류항목을 정리하였으며, 'Properties' 탭에서 온톨로지 구조에 따른 각 영역의 속성을 정의하였다. 처방명과 약재명, 증상명, 개별 조문 등 구체적인 관계값을 선택하고 정의해야 하는 목록은 모두 'Individuals' 탭에서 '인스턴스'로 입력하고 Fig. 4와 같이 각 인스턴스가 가지는 속성에 따라 관계값을 설정하였다.

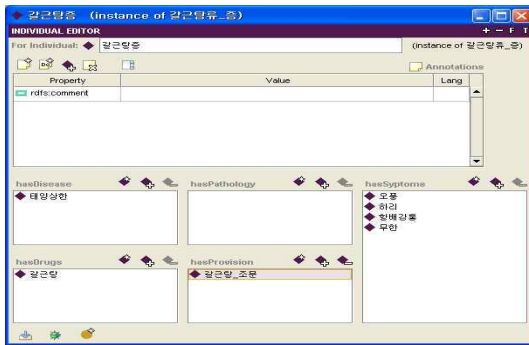


Fig. 4. 인스턴스와 속성 값 정의

은 계지탕 관련 조문을 보여주고 있다.

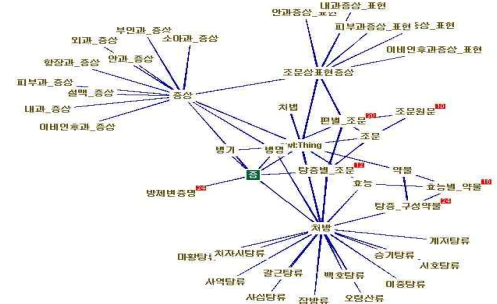


Fig. 6. 상한론 온톨로지 구조 - '증' 관련 영역

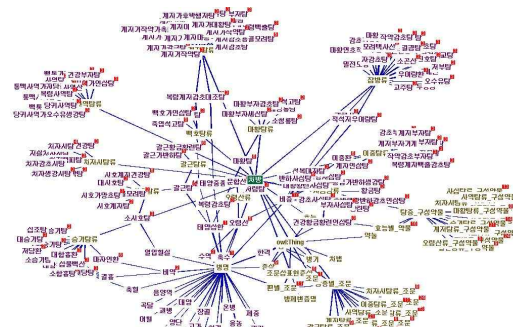


Fig. 7. 처방의 분류와 소속 처방

결 과

1. 온톨로지 구조

본 연구에서 설계하고 구축한 온톨로지 구조는 Fig. 5와 Fig. 6에서 보는 바와 같다. Fig. 5는 본 연구에서 다루는 영역과 각 영역 간의 속성을 표현한 것이며, Fig. 6은 '증'을 중심으로 관련 영역과의 관계를 표현한 것이다.

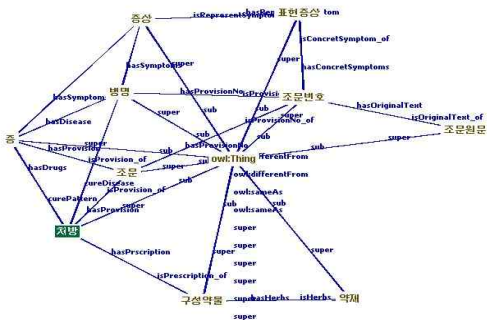


Fig. 5. 상한론 온톨로지 구조

2. 처방 온톨로지

각 처방은 분류 기준에 따라 소속 처방을 배치하고(Fig. 7), 개별 처방의 속성을 표현하기 위하여 개별 처방이 주치하는 증과 병, 구성약물 및 상한론 서적상의 관련 조문을 서로 연결하였다(Fig. 8). Fig. 9는 개별처방 중 계지탕의 속성을 표현할 수 있는 'Individual' 창을, Fig. 10은 계지탕의 구성약물을, 그리고 Fig. 11

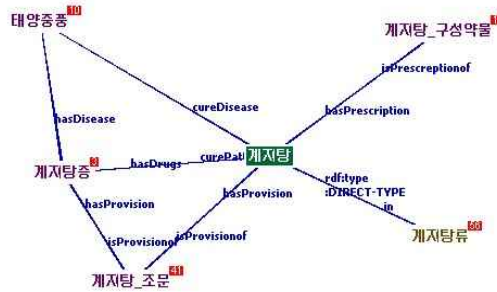


Fig. 8. 개별 처방의 속성

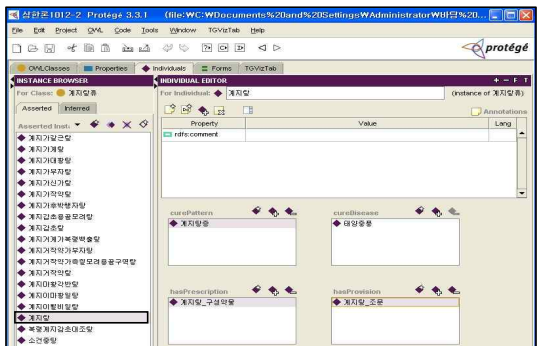


Fig. 9. 개별 처방의 속성 - 계지탕

3. 약제 온톨로지

약제 온톨로지는 본 연구에서 사용한 매개개념에 의한 '탕증

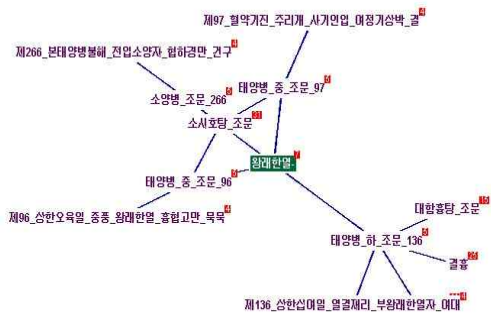


Fig. 17. 조문상 표현증상 중 '왕래한열' 관련 조문

5. 조문 온톨로지

조문은 '조문원문'과 '편별 조문' 및 '탕증별 조문'의 세 가지 데이터베이스를 구축하였다. 탕증 별 조문은 처방과 관련된 조문을 연결하기 위하여 설정한 매개개념으로 각 탕증과 관련 조문 번호를 연결하였고, 편별 조문은 상한론 각 편에 해당하는 조문 번호를 기재하고, 관련 조문원문과 조문에 표현된 증상들, 그리고 어느 처방의 조문인가를 Fig. 18과 Fig. 19처럼 설정하였다.

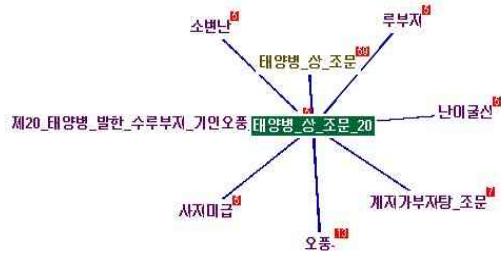


Fig. 18. 편별 조문 - 태양병 상 조문 20

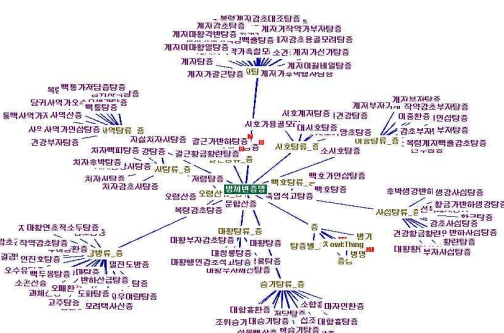


Fig. 19. 증 데이터베이스 - 증의 분류

6. 증 온톨로지

1) 증 데이터베이스

상한론에서의 증은 '처방중심의 증'이므로 본 연구에서도 처방 온톨로지와 동일한 명칭, 즉 '계지탕증'처럼 처방명에 '증'자를 붙여 사용하였으며, Fig. 20은 본 연구에서 다룬 증을 그래프로 표현한 것이다.

2) 증을 구성하는 증상의 정의

증을 구성하는 증상을 규정하기 위하여 우선 특정 증과 관

련된 조문을 탐색한 다음, 각 조문에서 표현된 '조문상 표현증상'의 '대표증상'을 찾아 이들의 조합으로 정의하였다. '대함홍탕'을 예로 들면, Fig. 20에서 보는 바와 같이 대함홍탕과 관련된 조문은 태양병 134, 135, 136, 137, 149번의 5개 조문이며, 각 조문에서 표현된 증상과 대표증상은 Fig. 21-24 및 Fig. 25에서 보는 바와 같고, 이들을 조합한 '대함홍탕증'은 Fig. 26과 같다.

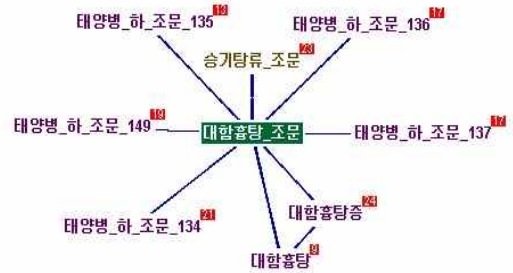


Fig. 20. 대함홍탕증 조문

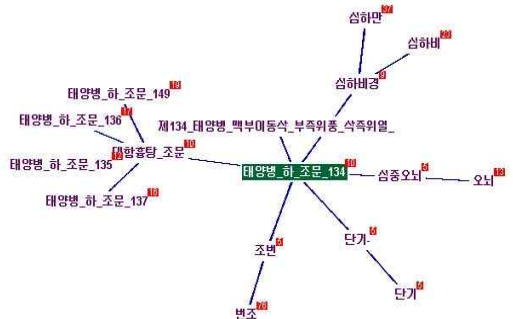


Fig. 21. 대함홍탕증 - 조문 134

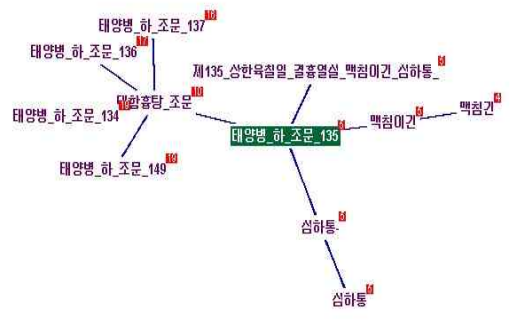


Fig. 22. 대함홍탕증 - 조문 135

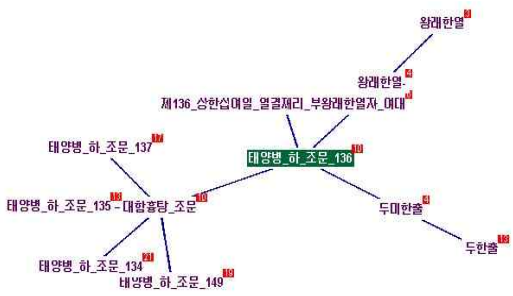


Fig. 23. 대함홍탕증 - 조문 136

8. 박경모, 임희숙, 박종현. Protege를 이용한 한의학의 구조화된 증상입력을 위한 온톨로지 개발. 동의생리병리학회지 17(5):1151-1156, 2003.
9. 신상우, 정길산, 박경모, 김선호, 박종현. 한의진단 Ontology 구축과 평가. 동의생리병리학회지 20(1):202-208, 2006.
10. 송미영. 온톨로지 기반 한의학 지능형 정보체계 연구(1차년도). 한국한의학연구원 보고서, 2009.
11. 송미영. 전통약재 정보의 효율적인 활용을 위한 온톨로지 구축. 한국한의학연구원 보고서, 2007.
12. 이재규, 최형립, 김현수, 서민수, 주석진, 지원철. 전문가시스템 원리와 개발. 법영사, 2002.
13. 이현실. 합성 온톨로지 기반의 한의학 처방 지식관리 시스템. 한국학술정보(주). 2006.
14. 임희숙. 온톨로지 기반 전자진료부 입력 인터페이스 설계 및 구현. 경희대학교 대학원 동서의료공학과 석사학위논문, 2004.
15. 張啓基, 王輝武, 郭自光. 傷寒論手冊. 과학기술문헌출판사 중경분사, 1983.
16. 전국한외과대학 본초학공동교재 편찬위원회. 본초학. 영림사, 2010.
17. Deborah, L. McGuinness. Frank van Harmelen. OWL Web Ontology Language Overview. W3C. 2004. (<http://www.w3.org/TR/owl-features/>).