

형식개념분석 기법을 이용한 임상의사결정지원시스템의 구축

강유경*, 황석형*, 김홍기**, 백승학**, 김동순*, 김응희*, 양경모*, 양성권*

*선문대학교 컴퓨터정보학부, **서울대학교 치과대학

e-mail:{aquamint99, shwang}@sunmoon.ac.kr,

hgkim@snu.ac.kr, drwhite@unitel.co.kr,

{ds8025, eunghee1, yjj0309, nrzbah}@gmail.com,

Development of a Clinical Decision Support System using Formal Concept Analysis

Yu-Kyung Kang*, Suk-Hyung Hwang*, Hong-Gee Kim**, Seung-Hak Baek**,

Dong-Soon Kim*, eung-hee Kim*, kyoung-mo Yang*, Sung-kwon Yang*

*Div. of Computer and Information Science, Sunmoon University,

**School of Dentistry, Seoul National University

요 약

방대한 양의 다양한 데이터들이 발생하는 의료분야에서는 임상데이터를 기반으로 보다 정확하고 효율적으로 현상을 분석/판단하여 의사가 환자진료 시 정확한 진단과 치료를 수행할 수 있도록 보조해주는 적절한 의사결정지원시스템이 요구되고 있다. 따라서, 이와 같은 요구를 충족시키기 위해서는 다종 다양한 데이터로부터 간결하면서도 효과적으로 개념들을 추출하고 구조화하여 개념계층구조로 표현할 수 있어야 하며, 실세계의 데이터에 대한 구조화와 요약의 제공과 필요한 정보를 수월하게 접근할 수 있어야 한다. 본 연구에서는, 도메인 내의 다양한 데이터들로부터 개념들을 추출하고, 개념들 사이의 상하위 관계를 파악하여 개념계층구조를 구축하기 위한 정형화된 데이터분석기법으로서 형식개념분석기법(FCM: Formal Concept Analysis)을 소개하고, 이를 치과 교정학 분야의 환자 임상데이터 분석기법(Cephalometric Analysis)에 융합한 형태의 임상의사결정지원시스템 개발 및 향후 연구과제 등에 관해 설명한다.

1. 서론

대부분의 조직들이 대량의 데이터를 가지고 있지만 그 데이터들을 효율적으로 잘 활용하지 못하고 있다. 그것은 데이터들을 구조적인 형태로 재구성하기 어렵고, 구조화 되지 않은 대량의 데이터 속에서 정확한 분석과 함축된 지식을 추출하기가 쉽지 않기 때문이다[1].

형식개념분석기법(FCA: Formal Concept Analysis)은 도메인 내의 다양한 데이터들로부터 객체(Object)와 속성(Attribute)들을 추출하고, 이들 사이의 관계를 파악하여 개념(Concept)을 생성하고 개념계층구조(Conceptual hierarchy)를 구축하기 위한 수학적 모델을 제공한다[2-4]. 개념계층구조를 토대로 제반 문제들에 대한 분석과 해결책을 모색하기 위한 응용수학분야의 분석기법인 FCA는 온톨로지공학, 세만틱 웹, 데이터 마이닝, 소프트웨어공학, 정보 검색 분야 뿐만 아니라 의학 및 바이오인포매틱스 등의 분야에 널리 적용되기 시작하고 있다[5-7].

의료분야에서는 매일 방대한 양의 다종다양한 데

이터들이 축적되고 있으며 환자진료 시 진단 및 치료를 위해 필요한 임상데이터들을 정확하게 분석하여 환자를 진료할 수 있도록 하기 위한 의사결정지원시스템이 요구되고 있다. 따라서 본 연구에서는 형식개념분석기법을 기반으로 임상데이터로부터 계층적 개념구조를 구축함으로써, 다량의 임상데이터들을 정확하게 분석하여 환자진료 시 의사의 진단 및 처방에 대한 의사결정을 지원할 수 있도록 임상 의사결정지원시스템(CDSS : Clinical Decision Support System)을 개발하였다.

본 논문은 다음과 같이 구성된다. 2장에서는 형식 개념분석기법의 기본적인 수학적 정의와 정형화된 계층적 개념구조의 구축방법에 대해 설명한다. 그리고 3장에서는 본 연구에서 개발된 임상의사결정지원 시스템에 대해서 소개하고, 4장에서는 결론과 향후 연구과제에 대해서 설명한다.

2. Formal Concept Analysis

2.1. One-valued context

1) 본 연구는 보건복지부 용어 표준화 지원도구 및 온톨로지 기반의 EHR 상호운용 기술개발 과제(과제번호:A05-0909-A80405-05N1-00050B)의 지원에 의해 이루어진 것임.

FCA는 개념격자라는 수학적 모델을 기반으로 하는 데이터분석기법의 일종으로서, 개념적인 데이터 분석과 knowledge processing분야의 제반문제들에 대한 수학적 해법을 제공하고 있다.

FCA의 기본이 되는 Formal context $K=(G, M, I)$ 는 객체들(Objects)의 집합 G 와 속성들(Attributes)의 집합 M , 그리고 G 와 M 사이의 이항관계 $I \subseteq G \times M$ 로 구성된다. 즉, 어떤 객체 g 가 속성 m 을 가지고 있을 경우, gIm 또는 $(g, m) \in I$ 로 나타내며, g 는 m 을 갖는다는 것을 의미한다. 표 1의 경우 7개의 음료수를 객체 G 로 음료수의 5가지 특징들을 속성 M 으로 하는 Formal Context를 나타내고 있다.

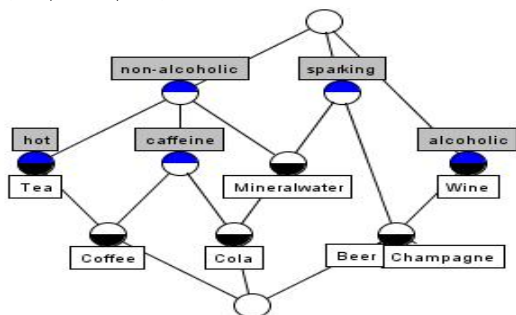
<표 1> 음료수에 대한 Formal Context

	non-alcoholic	hot	alcoholic	caffeine	sparkling
Tea	X	X			
Coffee	X	X		X	
Mineralwater	X				X
Wine			X		
Beer			X		X
Cola	X			X	X
Champagne			X		X

이때, Formal context $K=(G, M, I)$ 에 대하여, $O \subseteq G, A \subseteq M$ 일 때, $O' = A \wedge A' = O$ 를 만족하는 (O, A) 를 개념(formal concept)이라고 한다. 단, $O' := \{a \in M \mid \forall o \in O: (o, a) \in I\}$, $A' := \{o \in G \mid \forall a \in A: (o, a) \in I\}$.

즉, formal concept (O, A) 는 O 의 모든 객체들이 공통적으로 갖는 속성들의 집합이 A 와 같고, A 의 모든 속성들을 공통적으로 갖는 객체들의 집합이 O 와 같음을 의미한다. 또한, 임의의 개념 $(O_1, A_1), (O_2, A_2)$ 에 대하여, $O_1 \subseteq O_2 (\Leftrightarrow A_1 \supseteq A_2)$ 라면, (O_1, A_1) 은 (O_2, A_2) 의 상위개념(또는, (O_2, A_2) 는 (O_1, A_1) 의 하위개념)이며, $(O_1, A_1) \leq (O_2, A_2)$ 와 같이 표현한다.

Formal context $K=(G, M, I)$ 로부터 만들어진 모든 개념들과 그들 사이의 상위-하위개념관계 \leq 로 이루어진 계층구조를 개념격자(Concept Lattice 또는 Galois Lattice)라고 부른다. 개념격자는 Hasse Diagram을 사용하여 계층적인 개념구조형태로 가시화할 수 있다(그림 1 참조).



(그림 1) 음료수에 대한 개념격자

개념격자를 나타낸 Hasse Diagram에서는, 각 개념들과 이들 사이의 상하위관계가 링크에 의해 표시되며, 특히, 개념들 간의 링크에 의해 만들어지는 경로에 의해 상위개념으로부터 하위개념으로 속성들이

상속되며, 하위개념으로부터 상위개념으로 해당 객체들이 전파된다. 예를 들어, Tea는 고유속성으로 hot을 갖고 있으며, 상위개념들로부터 non-alcoholic 속성을 상속받는다. 한편 sparkling을 속성으로 갖는 객체로서는 Mineralwater, Cola, Beer, Champagne이 됨을 알 수 있다. 이와 같은 방법을 사용함으로써, 주어진 문제영역의 객체들과 이들이 갖는 속성들을 context형태로 파악하여, 개념을 추출하고 개념격자형태로 나타냄으로써, 도메인 내의 개념들을 분류하고 체계화 할 수 있는 계층적 개념구조를 수월하게 구축할 수 있다.

2.2. Many-valued context와 Scaling

FCA는 여러 가지 다양한 값을 가지는 속성들과 객체들, 그리고 객체와 속성 사이의 관계를 나타내는 데이터에도 적용될 수 있다. 예를 들어, "size"라는 속성은 "small", "medium" 또는 "large" 등과 같은 다양한 값들을 가질 수 있다. 이러한 속성들을 many-valued attributes라고 부르고 many-valued attributes를 포함한 context를 Many-valued context라고 하며 정형화된 정의는 다음과 같다.

[정의 1] Many-valued context $K=(G, M, W, I)$ 는 객체들(Objects)의 집합 G 와 속성들(Many-valued Attributes)의 집합 M , 속성의 값 W , 그리고 G 와 M 과 W 사이의 관계 $I \subseteq G \times M \times W$ 로 구성된다.■

즉, G 와 M 의 원소들은 각각 해당 context의 객체들과 각 객체들이 가질 수 있는 속성들, 그리고 그 속성의 값들을 나타낸다. 또한, 어떤 객체 g 가 속성 m 을 가지고 있고 그 속성의 값이 w 인 경우, $(g, m, w) \in I$ 또는 $m(g)=w$ 로 나타내며, 객체 g 는 w 값을 가지는 속성 m 을 갖는다는 것을 의미한다.

Many-valued context도 One-valued context와 같이 table로 나타낼 수 있으며, table의 각 셀에는 "X"표시 대신 해당 객체가 갖는 속성들의 값을 표시한다(표2 참조).

<표 2> 치과 교정학 분야에서 발생한 한 환자의 임상데이터에 대한 Many-valued Context

	Saddle angle	Articular angle	Gonial angle	Sum
Data1	133.32	145.88	134.47	413.68
Data2	135.55	148.19	135.42	419.17
Data3	135.27	148.81	134.62	418.71

이와 같은 Many-valued context로부터 개념들을 추출하고 개념격자를 구성하기 위해서는 특정한 규칙에 따라 Many-valued context를 One-valued context로 변환할 필요가 있다. 이와 같이 변환된 one-valued context를 derived context라고 부르며 이러한 변환과정을 conceptual scaling이라고 한다. conceptual scaling하기 위해서, Many-valued context의 각 속성들은 conceptual scale이라고 불리는 context를 토대로 One-valued context로 변환되며, Many-valued context의 각 속성들은 conceptual scale에 의해서 해석된다.

[정의 2] Many-valued context의 속성 m 에 대한 (conceptual) scale은 one-valued context $S_m := (G_m,$

M_m, I_m)로 정의된다. 단, $m(G) \subseteq G_m$. scale의 객체들은 scale values, 속성들은 scale attributes라고 부른다. ■

<표 3> Saddle angle 속성에 대한 scale context

$S_{\text{saddle angle}}$	-3 S.D.	-2 S.D.	-1 S.D.	N	+1 S.D.	+2 S.D.	+3 S.D.
Saddle angle < 109.49	X						
109.49 ≤ Saddle angle < 114.81		X					
114.81 ≤ Saddle angle < 120.13			X				
120.13 ≤ Saddle angle ≤ 130.77				X			
130.77 < Saddle angle ≤ 136.09					X		
136.09 < Saddle angle ≤ 141.41						X	
141.41 < Saddle angle							X

(Mean:125.45, S.D.(Standard Deviation):5.32, N:Normal)

<표 4> 표2에 대한 Scaling 적용 결과

	Saddle angle						
	-3S.D.	-2S.D.	-1S.D.	N	+1S.D.	+2S.D.	+3S.D.
Data1					X		
Data2					X		
Data3					X		

Many-valued context를 One-valued context로 변환하기 위한 scale들은 이외에도 다수 존재할 수 있으며, scaling 하고자 하는 속성의 특성에 따라 scale를 선택하여 scaling하면 된다. 예를 들어, 표 2와 같은 Many-valued context의 각 속성들에 대하여 표3의 Nominal scale을 적용하여 손실된 정보 없이 표 4와 같이 one-valued context로 변환할 수 있다.

위의 각 정의들을 토대로 본 연구에서는 치과 교정학 분야에서 발생된 임상데이터(many-valued context)를 scaling하여 one-valued context로 변환한 후 치과 교정학 분야의 환자 임상데이터 분석을 위한 Cephalometric analysis 기법을 융합한 형태로 개념격자를 표현하였다(그림 2 참조).

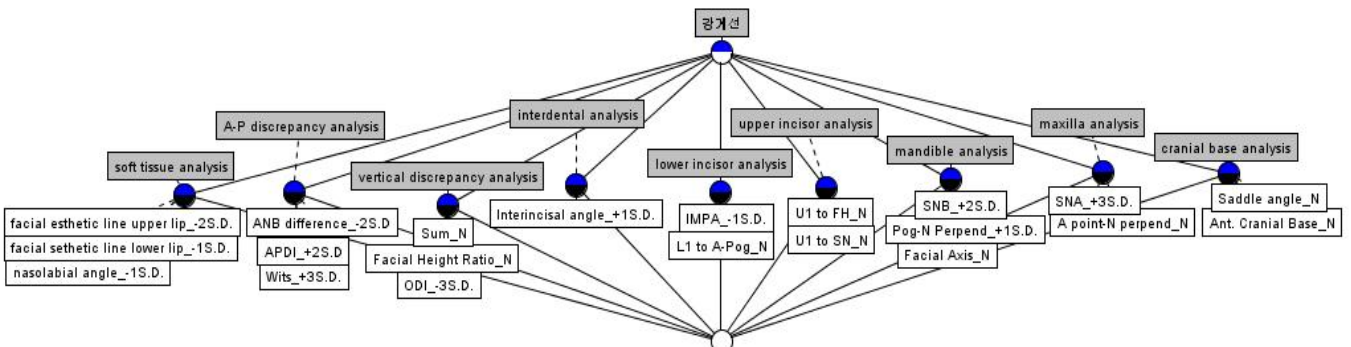
3. Clinical Decision Support System

임상의사결정지원시스템(CDSS: Clinical Decision Support System)은 의료 분야에 있어서 의사결정을 지원해 주는 시스템으로서, 의사가 환자의 임상 데이터를 분석하여 의료분야의 전문지식을 토대로 진

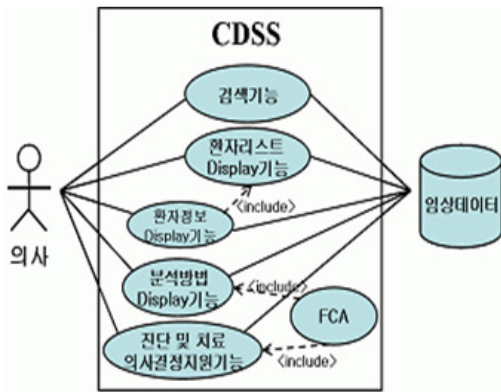
단 및 치료에 대한 의사결정을 할 때 사용된다. 임상 의사결정지원시스템은 비정상적인 결과에 대해 경고를 하는 알림시스템, 실제 진료에 있어서 환자의 이전 의무기록을 분석하여 의사결정에 도움을 주도록 구축된 진단시스템, 어떤 일정에 대해 의사가 있고 넘어가는 일이 없도록 상기시켜주는 신호시스템 등이 있다[8].

본 논문에서 개발된 임상 의사결정지원시스템은 치과 교정학 분야에서 사용하기 위해 개발된 시스템으로서, 환자의 임상데이터를 분석하여 진단 및 치료에 대한 의사의 의사결정을 지원하는 진단시스템으로 그 구성은 다음과 같다(그림3 참조).

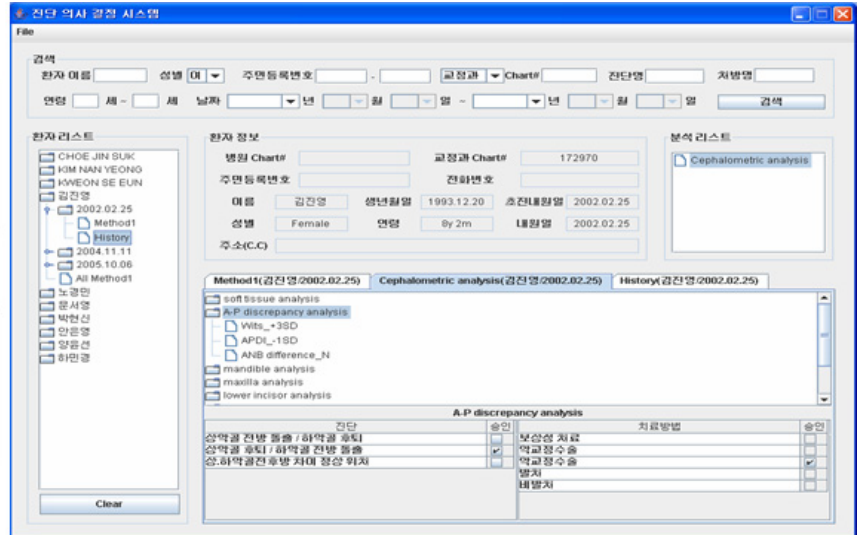
- 1) 검색기능** : 환자를 검색하는 기능으로서 환자의 기본 정보인 환자이름, 성별, 주민등록번호, 병원 또는 교정과 차트번호로 검색가능하고, 이미 CDSS를 사용하여 진단과 처방이 내려진 환자에 대해서 진단명과 처방명으로 검색할 수 있다. 또한 “10세~15세”와 같이 연령별로 환자를 검색할 수 있으며 날짜별로도 검색가능하다. 뿐만 아니라 “10세~15세 사이의 3급 전치부 개방교합(진단명)”과 같이 모든 검색 항목들은 연결해서 검색할 수 있다.
- 2) 환자리스트 Display기능** : 검색 결과를 보여주는 부분으로 환자 이름으로 표현되며 그 하위에는 진찰받은 날짜별로 Method 1파일과 진단 및 처방에 대한 History파일이 있다.
- 3) 환자정보 Display기능** : 진단 및 처방을 하기 위해 환자리스트에서 선택된 환자의 기본 정보 및 주소(cc: chief complain)를 보여준다. 주소를 보여줌으로써 의사가 진단과 처방에 대한 의사결정을 할 때 환자의 치료 목적에 초점을 맞추어 결정할 수 있다.
- 4) 분석방법 Display기능** : 환자의 임상데이터를 분석하기 위한 분석기법을 나타낸다. Cephalometric analysis [9,10]는 치과 교정학 분야에서 환자의 임상데이터를 분석하기 위한 가장 기본적인 분석 기법이고, 이 이외에도 여러 종류의 분석 기법들이 있으며 향후 추가 할 예정이다.
- 5) 진단 및 치료 의사결정지원기능** : 환자의 임상데이터에 형식개념분석기법을 적용하여, 환자의 어떤 계측치가 한국인 표준 계측치와 얼마만큼의 차이를 나타내는지 7단계의 S.D(Standard Deviation)



(그림 2) 한 명의 환자에 대한 임상데이터 분석결과와 Cephalometric analysis를 융합한 개념격자



(a) CDSS의 use case diagram



(b) CDSS의 Screenshot

(그림 3) CDSS의 use case diagram과 Screenshot

으로 표현한다(표3 참조). 각 계측항목에 대한 분석결과에 따라서 진단과 치료방법을 추천해 준다. 의사는 계측항목의 분석결과와 추천된 진단 및 치료방법의 타당성을 판단하여 진단과 치료방법을 선택하게 된다. 선택된 진단과 치료방법은 History파일에 저장되어 최종 Report 형식으로 저장 및 출력이 가능하다.

지금까지는 치과 교정학 분야에서 의사가 진단을 하기 위해서는 약 53개의 계측항목들의 계측치를 분석하고 관련된 계측항목들을 찾아서 비교하여 그 결과를 토대로 의사의 임상 경험과 지식에 따라 진단을 하였다. 그러나, 본 연구에서 개발된 임상 의사결정 지원시스템은 형식개념분석기법을 기반으로 53개의 계측치 분석 및 관련 있는 계측항목들을 자동적으로 분류해주고, 그 분석 결과를 토대로 진단과 치료방법을 추천해준다.

CDSS는 환자의 임상데이터를 분석하는데 드는 시간을 줄일 수 있으며, 잘못된 판단으로 발생하는 의료사고를 줄이는데 도움이 될 수 있다. 뿐만 아니라 현재 진단하고자 하는 환자와 가장 유사한 환자를 검색 할 수 있으며, 유사한 환자의 계측치 비교 및 진단, 치료방법, 치료결과 등을 비교하여 좀 더 정확한 진단과 효율적인 치료방법을 선택할 수 있다.

4. 결론 및 향후 연구 과제

환자를 진료하는데 있어서 의사는 진단과 치료에 관한 많은 의사결정을 해야 한다. 본 논문에서는 형식개념분석기법을 기반으로 환자의 임상데이터를 자동적으로 분석하여 환자진료 시 보다 수월하게 진단 및 처방을 할 수 있도록 지원하는 임상 의사결정 지원 시스템(CDSS)을 개발하였다. CDSS를 사용함으로써, 환자 임상데이터를 분석하는데 걸리는 시간과 노력을 줄일 수 있으며, 유사한 환자를 비교하여 정확한 진단과 효율적인 처방을 할 수 있다. 본 연구에서 개발된 CDSS는 OO대학병원 치과 교정학 분야에서 활용할 예정이며, 현재는 Cephalometric

analysis기법만을 사용하여 임상데이터를 분석하고 있으나, 향후에는 다양한 측면으로 환자 임상데이터를 분석하여 정확한 진단과 처방을 할 수 있도록 여러 종류의 치의학분야의 임상자료 분석기법들을 추가 할 예정이다.

형식개념분석기법을 기반으로 하는 향후 연구과제로서는, 치과 교정학분야 뿐만 아니라 여러 의료 분야에서 활용 할 수 있는 임상 의사결정 지원시스템을 구축하기 위한 제반연구와 대량의 임상데이터를 좀 더 의미적으로 분석하기 위한 Scaling 기법에 대해 연구할 필요가 있다.

참고문헌

- [1] Jeong-Jin Park, Yung-Moon Chae, Suk-II Kim, Jeong-Eun Kim, Gon-hee Kim, "효율적 암 분류를 위한 텍스트마이닝 기반의 의학적 의사결정지원시스템 개발", 대한의료정보학회지, 11권, 보완본2호, pp.146-150. 2005.
- [2] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundations, Springer, 1999.
- [3] C. Carpineto, G. Romano, Concept Data Analysis: Theory and Applications, Wiley, September, 2004.
- [4] B.A. Davey, H.A. Priestley, Introduction to Lattices and Order, Cambridge University Press, 2002.
- [5] S. Hwang, H.G. Kim, H.S. Yang, "A FCA-based Ontology Construction for the Design of Class Hierarchy", ICCSA2005, LNCS3481, Springer, 2005.
- [6] M. Obitko, V. Sansel, J. Smid, "Ontology Design with Formal Concept Analysis", Proc. of the CLA2004 Ostrava, Czech Republic, Sept. 23-24, 2004
- [7] 강유경, 황석형, 최성희, 최희철, 김동순, 김흥기, 김명기, "임상자료로부터 개념계층구조 구축을 위한 형식개념분석 도구의 개발 및 적용", 대한의료정보학회지, 제11권, 보완본2호, pp. 71-76, 2005.
- [8] 범희승, 박성희, 최진욱, 김준배, "임상 의사결정 지원 시스템의 약제부작용 감소 효과에 대한 메타분석", 대한의료정보학회지 8(3), 55-60, 2002.
- [9] 양원식, 김태우, 치과교정진단 및 응용, 지성출판사, 2001.
- [10] 전국치과대학교정학교수협의회, 치과교정학, 지성출판사, 1998.