

ITS : 지능적 Tissue Mineral Analysis 의료 정보 시스템

ITS: Intelligent Tissue Mineral Analysis Medical Information System

조영임

Young Im Cho

Dept. of Computer Science, The University of Suwon

요 약

현재 국내에서는 TMA(Tissue Mineral Analysis) 결과를 독자적이며 전문적으로 해석할 수 있는 의료 정보 데이터베이스가 없을 뿐 아니라, TMA 관련 데이터베이스가 있다 하더라도 의료기관에서 사용하기 어려운 매우 낮은 수준이므로 양질의 의료서비스를 제공하기 어려운 실정이다. 또한 국내에서 주로 사용되는 TMA 방법은 미국에 의뢰한 후 보내온 분석결과에 의존하게 되는데, 이때의 결과는 서구식 생활패턴에서 비롯된 데이터베이스에 의해 분석된 것이므로 동양인의 경우 특히 검사결과의 신뢰성 문제가 발생하게 된다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해 국내 임상자료를 바탕으로 TMA 관련 국내 최초 지능적 의료정보 시스템(ITS: Intelligent TMA Information System)을 개발하였다. ITS는 TMA 자료를 다단계 통계분석 방법에 의한 결정트리 분류기를 이용하여 분류하고 다중 퍼지 규칙베이스를 구축하여 복잡한 자료를 추론하도록 구축하였다.

Abstract

There are some problems in TMA. There are no databases in Korea which can be independently and specially analyzed the TMA results. Even there are some medical databases, some of them are low level databases which are related to TMA, so they can not serve medical services to patients as well as doctors. Moreover, TMA results are based on the database of american health and mineral standards, it is possibly mislead oriental, especially korean, mineral standards. The purposes of this paper is to develop the first Intelligent TMA Information System(ITS) which makes clear the problems mentioned earlier. ITS can analyze TMA data with multiple stage decision tree classifier. It is also constructed with multiple fuzzy rule base and hence analyze the complex data from Korean database by fuzzy inference methods.

Key Words : TMA, 의료정보 데이터베이스, 다중 퍼지규칙베이스

1. 서 론

인체의 영양상태를 스크리닝하는 방법은 크게 3가지가 있는데 혈액검사, 요 검사, TMA(Tissue mineral Analysis)에 의한 검사 등이 있다. 혈액검사는 세포외액의 영양상태를 나타내고 있어서, 세포내의 영양상태를 정확하게 반영하지 못하며 일과성 변화가 매우 심하다는 단점을 갖는다. 또한 요 검사는 인체의 배설상태를 나타내고 있어서 세포내의 대사를 정확하게 반영하지 못하는 단점을 갖는다. 그러나 TMA는 세포내의 대사상태를 가장 정확하게 반영하고 있어서 세계 10대 사망원인을 정확하게 측정하고 있으며 이미 세계 48개국에서 도입되어 사용되는 보편적인 방법이다[1,2]. TMA에 관해서는 카톨릭 대학병원, 분당 차병원, 고려대학교 병원, 삼성제일병원, 연세대학교 세브란스 병원[3] 피부과 및 50여개 종합병원과 지방 350여 병원에서 전문 의료진들이 TMA 및 임상영양에 대한 필요성을 인식하여 임상 중이

므로 데이터베이스 구축에 관한 연구의 필요성을 인식하고 있다.

일반적인 TMA 분석과정은 먼저, 환자의 모발 약 5그램을 채취하여 TMA 분석기기에 넣어 환자 개개인에 대해 38종 이상의 미네랄 분포량을 얻는다. 이로부터 환자의 현재 영양 상태 및 인체의 길항작용 등을 설명하게 되며, 예방 및 치료할 수 있는 다량의 영양자료와 식단을 데이터베이스로부터 분석하여 환자와 의사들에게 의학진단정보로써 제공한다. 그러나 인체 상태를 평가하는 TMA자료는 매우 복잡할 뿐 아니라 38종의 미네랄이 다양한 수치범위(-0.00001 ~ +60000)를 갖고 있어서 통계적으로 다양한 분석 및 추론기법이 필요하다.

현재 국내 TMA 분석방법은 분석결과를 해석할 수 있는 의학 전문 프로그램이 개발되어 있지 않고, 일부 국내에서 기구축한 TMA관련 데이터베이스가 매우 낮은 수준이어서 의료기관에서 충분한 의료서비스를 제공하기 어렵다. 그러므로 현재 국내에서는 대부분의 TMA 분석을 미국에 의뢰하게 되는데 이 결과가 서구식 생활패턴에서 비롯된 데이터베이스로부터 나온 것이므로 한국인에 대해서는 보다 분석적

접수일자 : 2004년 6월 26일

완료일자 : 2005년 2월 21일

이고 한국적인 TMA 결과분석이 필요하다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해 TMA 분석을 위한 알고리즘을 개발하고 TMA 분석결과에 대한 데이터베이스를 구축하여 지능적 의료 정보 시스템(ITS: Intelligent TMA Information System)을 구축하고자 한다. 이것을 위해 국내 미네랄 관련 임상자료를 이용하여 통계분석적 오류와 데이터베이스 구축 및 추론상의 오류를 최소화하기 위해 다중 퍼지규칙 베이스를 구축하였고 다단계 통계분류 방법을 이용하였다.

본 논문의 구성은 다음과 같다. 2장에서는 TMA에 대해 설명하고, 3장에서는 ITS의 주요 기능을 설명하며, 4장에서는 ITS의 구현 및 성능분석을 하고, 5장에서 결론을 맺고자 한다.

2. TMA(Tissue Mineral Analysis)

TMA란 머리카락 속에 있는 30여 가지의 미네랄과 8가지의 중금속의 양과 중요 미네랄 비율을 분석하여 체내에 과잉, 결핍 및 불균형 상태를 평가하고 대사타입(metabolic type)을 제시하여 건강을 유지하는 방향을 제시하는 임상 영양학 및 독성학 모발조직 검사방법을 말한다. 또한 호르몬의 세포수준 활성을 제시(갑상선 및 부신)하며 영양, 신경 및 호르몬의 불균형을 통해 질환발생 가능성을 설명한다[1].

TMA의 특징은 모발 표본은 비침해 방법으로 채취되므로 혈액, 소변, 기타 다른 조직보다 빠르고 쉽게 수집할 수 있고, 다른 방법의 미네랄 분석보다 비용 대비 효율이 우수하다. 모발은 혈액과는 달리 미량원소의 농도에 민감한 영향을 미치는 항상성 기전에 덜 민감하며 미네랄의 저류 또는 손실과 같은 장기간의 편차들은 혈액보다 모발조직에서 쉽게 검출되고, 대부분의 모발조직 미네랄 농도는 혈액이나 다른 조직에서 발견되는 미네랄보다 농도가 매우 높다. 모발조직은 과거 뿐만 아니라 현재의 미량원소의 수치기록을 제공하며, 혈청 뿐만 아니라 외부원인으로부터 모발로 들어가는 물질의 정보를 제공하고 독성 중금속 수치를 평가하는데 유용하다는 점이다.

혈액이나 혈장에는 각종 미네랄이 함유되어 있으나 이들이 인체 미네랄 상태를 완전하게 표시할 수 없는데, 그 이유는 미네랄은 항상성의 일환으로 미네랄의 혈장 수치는 조직 미네랄에 의해 유지되고 있으나 혈장농도는 사람의 기온이나 혈액채취시간 또는 표본 채취적의 음식 섭취에 따라 변동이 다양하기 때문이다. 일반적으로 혈액은 인체의 저장 부위로부터 또는 저장부위의 미네랄 이동의 좋은 지표가 된다. 인체에 미네랄이 과다하게 축적되면 종종 이들 미네랄이 혈액에서 제거되어 조직으로 축적되기 때문에 혈청에서 검출되지 않을 수도 있다고 한다. 따라서 혈액 측정 방법보다는 TMA에 의한 측정 방법이 더 정확하다고 할 수 있다.

3. ITS (Intelligent TMA Information System)

본 논문에서 개발한 TMA 분석을 위한 지능적 의료 정보 시스템은 다음 그림 1과 같이 TMA 자료의 통계적 분석 부분과 사용자 중심의 정보검색이 가능한 데이터베이스 구축부분으로 구성되어 있다. 통계적 분석이 완료된 후 TMA 데이터베이스를 구축하고, 임상자료를 통해 시스템을 검증한다.

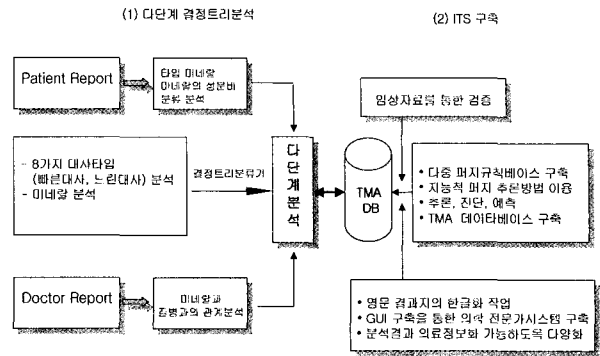


그림 1. ITS 개요

지능적 추론의 배경이 되는 대사타입의 분류 및 미네랄의 분석은 통계적 분류 및 분석 방법 중에서 결정트리분류기(Decision Tree Classifier)를 응용하여 TMA 자료의 대분류에 해당하는 인체의 8가지 대사타입(빠른 대사 4타입, 느린 대사 4타입)을 분류한다. 결정트리분류기는 두 종류이상의 결정을 적절한 순서에 따라 여러 번 반복하여 입력패턴의 클래스를 결정해야 할 때 매우 유용한 방법이다[4,5,6].

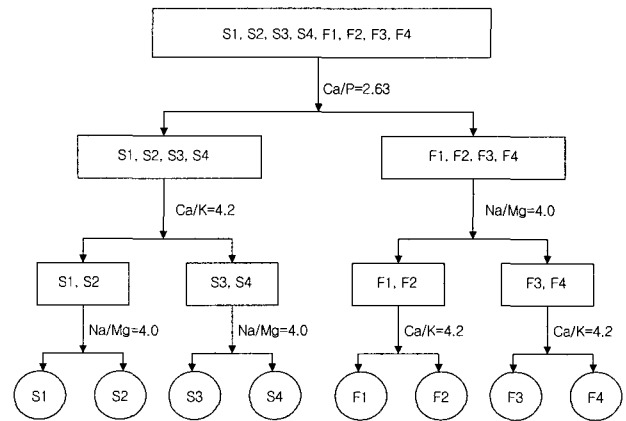


그림 2. 결정트리에 의한 TMA 타입 분류

그림 2는 결정트리분류기가 패턴집합을 분류한 것인데, 분류과정은 다음과 같다. 먼저 루트노드에서 미네랄 비 $Ca/P(=2.63)$ 를 이용하여 느린 대사타입(Slow Metabolic Type)의 패턴집합 $\{S_1, S_2, S_3, S_4\}$ 와 빠른 대사타입(Fast Metabolic Type) 패턴집합 $\{F_1, F_2, F_3, F_4\}$ 의 서브패턴집합으로 분류하고, 다시 내부노드에서 $Ca/K(=4.2)$ 를 이용하여 1차 분류된 서브패턴 집합 $\{S_1, S_2, S_3, S_4\}$ 를 패턴 $\{S_1, S_2\}$ 와 $\{S_3, S_4\}$ 로 분류하고 $Na/Mg(=4.0)$ 를 이용하여 1차 분류된 서브패턴 집합 $\{F_1, F_2, F_3, F_4\}$ 를 $\{F_1, F_2\}$ 와 $\{F_3, F_4\}$ 로 분류한다. 마지막 단계에서 $Na/Mg(=4.0)$ 를 이용하여 2차 분류된 패턴집합 $\{S_1, S_2\}$ 와 $\{S_3, S_4\}$ 를 S_1 과 S_2 그리고 S_3 과 S_4 로 최종 분류하고, 같은 방법으로 $Ca/K(=4.2)$ 를 이용하여 2차 분류된 패턴집합 $\{F_1, F_2\}$ 와 $\{F_3, F_4\}$ 를 F_1 과 F_2 그리고 F_3 와 F_4 로 최종 분류하여 결정트리를 구성한다. 대사타입의 분류는 특성변수에 대한 클래스별 분포를 이용해도 같은 결과를 보여 준다.

다음은 특성변수를 계급적으로 적용하여 클래스별 분포를 이용한 대사타입의 분류를 보여준 것이다. 우선 Ca/P 클래스

분포에 의해 대사타입을 $\{S_1, S_2, S_3, S_4\}$ 와 $\{F_1, F_2, F_3, F_4\}$ 의 서브패턴집합으로 분류한 것이 그림 3이다.

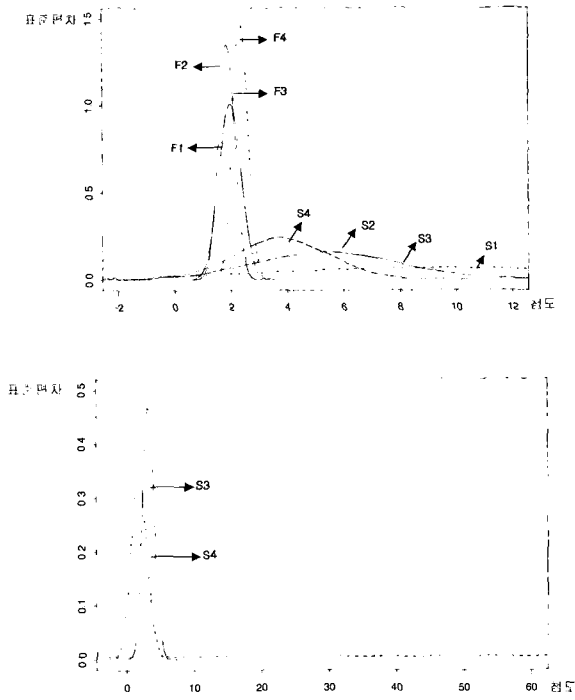


그림 3. F타입의 Ca/P과 Ca/K 클래스 분포를 이용한 대사타입의 분류

그림 3 왼쪽의 4개 분포는 점도가 높고 표준편차가 작은 대사타입 F의 Ca/P분포로 $F_1 \sim (\hat{\mu}=1.97, \hat{\sigma}=0.38)$, $F_2 \sim (\hat{\mu}=1.85, \hat{\sigma}=0.28)$, $F_3 \sim (\hat{\mu}=2.01, \hat{\sigma}=0.32)$, $F_4 \sim (\hat{\mu}=2.34, \hat{\sigma}=0.26)$ 를 보여주고 있으며 아래쪽으로 누운 형태의 4개 분포는 점도가 낮고 표준편차가 큰 대사타입 S의 Ca/P분포로 $S_1 \sim (\hat{\mu}=9.31, \hat{\sigma}=6.65)$, $S_2 \sim (\hat{\mu}=5.35, \hat{\sigma}=3.45)$, $S_3 \sim (\hat{\mu}=5.75, \hat{\sigma}=3.60)$, $S_4 \sim (\hat{\mu}=3.79, \hat{\sigma}=1.64)$ 을 보여주고 있어 Ca/P의 분포에 의해 대사타입 F와 S가 분류됨을 볼 수 있다. Ca/K의 분포는 점도가 높고 표준편차가 작은 $S_4 \sim (\hat{\mu}=1.78, \hat{\sigma}=1.15)$, $S_3 \sim (\hat{\mu}=3.16, \hat{\sigma}=0.81)$ 가 왼쪽에 위치하고 아래쪽에 점도가 낮고 표준편차가 큰 $S_2 \sim (\hat{\mu}=11.7, \hat{\sigma}=11.03)$, $S_1 \sim (\hat{\mu}=54.79, \hat{\sigma}=71.8)$ 로 분류된다.

Ca/K의 분포에 의해 분류된 S_1 과 S_2 , S_3 와 S_4 를 Na/Mg의 분포에 의해 분류하면 다음 그림 4와 같다.

분류방법의 요점은, TMA에서 가장 기본이 되는 대사타입의 분류 후에 각 대사타입 별 영양미네랄과 독성미네랄의 영양상태와의 관계는 상관분석을 통해 상관계수가 0.3보다 큰 미네랄들을 대상으로 분석한 점이다.

분석된 통계를 바탕으로 본 논문에서는 불확실한 다량의 정보로부터 지능적 추론이 가능한 데이터베이스 구축을 위해 다중 퍼지 규칙베이스를 구축한다. 다중 퍼지 규칙베이스는 TMA 자료와 같이 다양한 미네랄에 의해 의사결정을 해야 하는 시스템인 경우 규칙베이스가 매우 복잡해 질 수 있기

때문에 규칙베이스를 분산시켜서 연관성을 찾아 실시간으로 규칙들을 검색하고 추론하는 방법이 필요하다. 이것을 위해 본 논문에서는 규칙들의 결론부와 다른 규칙의 조건부와의 관계를 트리형태로 구성하여 다단계로 구성함으로써 기존의 퍼지 규칙베이스보다 복잡한 규칙을 표현하기에 적합하도록 하는 방법을 제안한다. 또한 퍼지 이론을 본 논문에서 불확실성 처리를 위해 적용한 이유는 베이시안 이론 등 불확실성 처리방법들 중[7] 자료처리 및 제어시스템에 적합하기 때문이다.

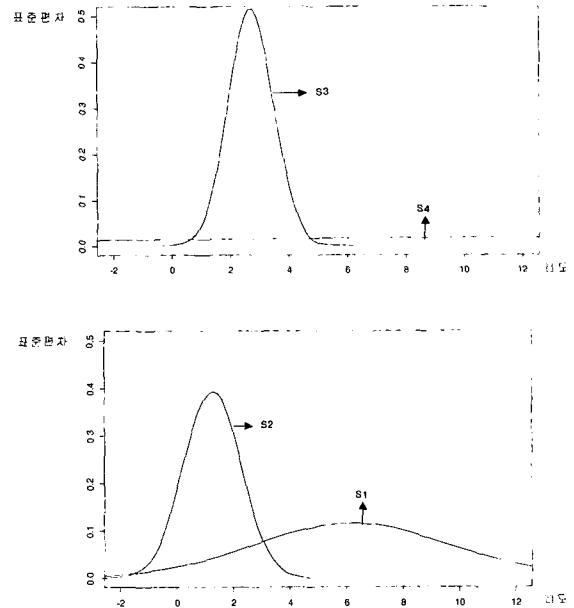


그림 4. S_1 , S_2 와 S_3 , S_4 타입의 Na/Mg 클래스 분포를 이용한 대사타입의 분류

다중 퍼지규칙 베이스 구축의 제 1단계는 빠른 대사와 느린 대사를 결정짓는 미네랄들의 상관관계 분석으로부터 대사타입에 따른 퍼지 데이터베이스를 구축하는 단계이다. 이것을 위해 본 논문에서는 영양미네랄 30종과 독성 미네랄 8종류 등 38종 미네랄과 영양상태와의 관계, 영양미네랄들의 비, 독성미네랄들의 비와 영양상태와의 관계, 그리고 TMA를 통한 각종질환과 미네랄과의 경향에 관해 분석하였다.

문장 분석시 영향을 미치는 주요 미네랄 비율을 나타내는 언어적 변수인 "높다(high)" "낮다(low)" "적당(acceptable)" 등은 퍼지 집합[8]으로 표현하였다.

이와 같은 7종의 미네랄의 비 뿐 아니라 각 미네랄(Ca, Na, Mg 등 영양미네랄 15종, Hg, Pb 등 독성미네랄 8종, Bi, Ge 등 기타 미네랄 15종) 등에 대해서도 위의 그림과 같이 퍼지 집합으로 표현하였다.

다중 퍼지규칙 베이스 구축을 위한 다음 단계는 앞선 대분류 과정을 통해 구축한 대사타입에 따른 퍼지 데이터베이스를 세분화하고 또한 내분비선과 미네랄과의 상관관계에 대한 데이터베이스를 구축하는 과정이다. 예를 들면, Ca/P, Ca/Mg, Ca/Na, Ca/K, Ca/Fe가 증가하면 Parathyroid가 증가하게 되고, 반대로 Fe/Cu가 증가하면 Parathyroid가 감소한다는 것과 같은 임상 실험결과를 분석하였다.

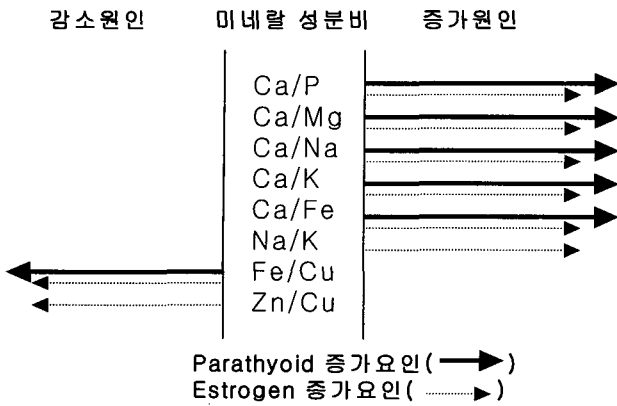


그림 5. 미네랄 비율과 내분비선과의 증감관계

이상의 방법을 통해 구축된 다중 퍼지 규칙베이스는 다음의 예와 같이 퍼지 규칙의 조건부에 조건명제들의 수를 다음과 같이 3개 이하로 제한시켜 데이터베이스에서의 추론 방법을 효율화 했으며 계층적 구조를 이용하여 연관된 다음 규칙과의 관계로부터 규칙의 조건부가 추론되도록 하였다. 여기서 데이터베이스에는 Zn/Cu, Na/K 값 등이 멤버쉽 함수로 표현하여 처리하였다.

if Na/K가 낮고(low) Zn/Cu가 높고(high) Fe/Cu가 높으면(high)
then Progesteron을 증가시킨다.

본 논문에서는 다단계 통계분석방법을 사용하여 분류한 자료로부터 구축한 다중 퍼지 규칙베이스에서 퍼지 추론을 효율적으로 하기 위해 분산 목표트리 방식인 TAEMS(Task Analysis Environment Modeling and Simulation)[9]를 적용하였다. TAEMS는 태스크를 분류한 후 분산 목표트리를 구성하여, 서로 다른 목표를 갖는 시스템의 분산목표 수행을 위해 시스템의 목표나 자료들 간의 상호 의존성을 나타냄으로써 효율성을 높이기 위한 분산 인공지능적인 방법이다. TAEMS에서 제일 위쪽의 루트 태스크는 작은 서브 태스크로 분류할 수 있고 서브 태스크는 더 이상 분해할 수 없는 메소드(method)로 분해된다. 가장 밑에 있는 leaf node는 메소드의 역할을 하는데 각 에이전트가 실질적으로 수행하는 행동을 나타내는 것이다.

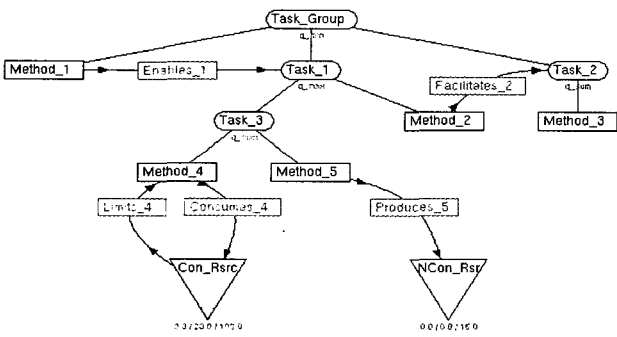


그림 6. TAEMS 태스크 구조

다음 그림 7은 다중 퍼지 규칙베이스에서 규칙들간의 상호의존성을 TAEMS 방식을 적용해서 나타낸 그림이다. 본 논문에서는 규칙의 조건부에는 미네랄의 비율을 이용하여 구축

하였고, 결론부에는 대사타입 및 규칙의 조건부에 입력되는 미네랄의 비율을 이용하였다. 본 논문에서 현재까지 구축한 규칙은 약 1000개 정도이다. 퍼지 규칙베이스로부터 추론되는 알고리즘을 설명하면, 규칙의 조건부와 입력사실이 일치하고 연관된 규칙을 찾아서 그것으로부터 규칙들간의 상호 의존성을 고려해서 생성된 규칙들로부터 값을 유도한다.

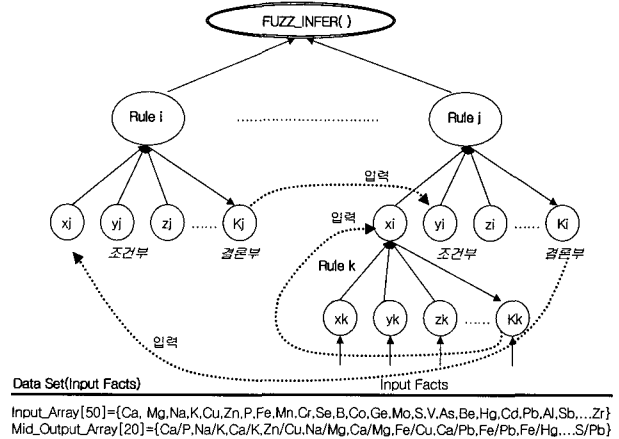


그림 7. 다중 퍼지 규칙베이스에서 규칙들 간의 상호의존관계

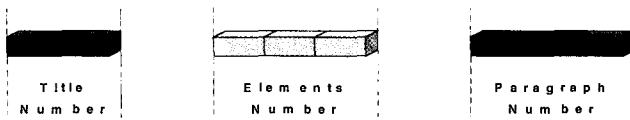
퍼지 규칙베이스에 있는 각 규칙들은 위의 그림과 같이 상호의존관계를 갖는다. 즉, 하나의 규칙을 구성하는 조건부, 결론부 퍼지명제들은 다른 규칙의 조건부에 값을 전달하는 상호의존관계로 구축되어 있다. 이렇게 하는 이유는 TMA 미네랄 자료가 단순히 몇 개의 미네랄에 의해 분류되지 않고 매우 복잡한 연관관계를 통해 분석되고 측정되기 때문이다.

이 구조에 따라 다중 퍼지 규칙베이스에서 10개의 상위레벨로부터 파생된 인체의 대사타입에 따라 분류·구축된 약 1000개의 규칙들의 각 조건부들을 찾아가는 과정을 예를 들어 나타내면 다음 그림 9와 같고 ID 작성원칙은 그림 8(a)와 같다. Title number는 Title에 대한 순서이며, Element number는 Title에 따른 각 element 순서이며, Paragraph number는 각 paragraph에 따른 문장의 순서이다. 예를 들어, ID가 '13002368'인 영어 문장은 출력해야 할 대분류 'Nutrient Mineral Level' 중 두 번째 미네랄인 'Mg'에 관한 368번째 문장을 말한다. 이 ID인 문장이 추론되기 위해서는 그림 9에서와 같은 3개의 퍼지 규칙이 연관되어 만족되어야 한다. 즉, 'THEN'이하의 결론부인 'S1', 'Ca/K is low', 'Na/Mg is high'가 각각 Rule #2의 조건부에 입력되어 TAEMS방식의 다단계 스케줄링 과정을 거쳐서 해당 ID 문장을 추론하게 된다. 각 규칙의 결론부가 상위레벨의 규칙의 각 조건부 문장들과 일치되는 것들을 찾아가면서 최종적으로 결론부의 문장 ID를 찾게 된다.

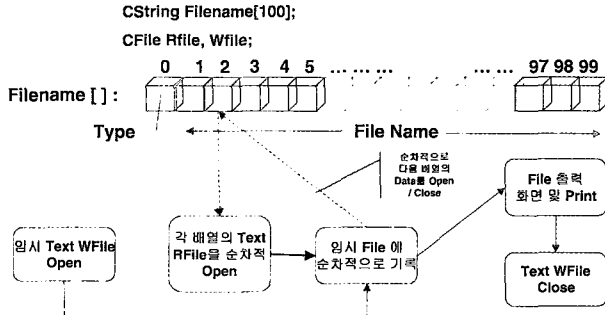
그림 8(b)는 ID로부터 다음 그림 9를 이용하여 문장을 하나하나씩 찾아내어 조합해나가는 과정을 나타낸 그림이다. 이로부터 추론된 해당 ID인 문장은 다음과 같다.

【ID: 13002368인 영어문장】

"Magnesium deficiency has been shown to be associated with decreased antibody production. Published studies have revealed that the lymphocytes, which are the body's defence against foreign invaders, are inhibited when there is a deficiency of Magnesium"



(a) ID 작성규칙



(b) ID로부터 문장조합과정

그림 8. ID 작성규칙

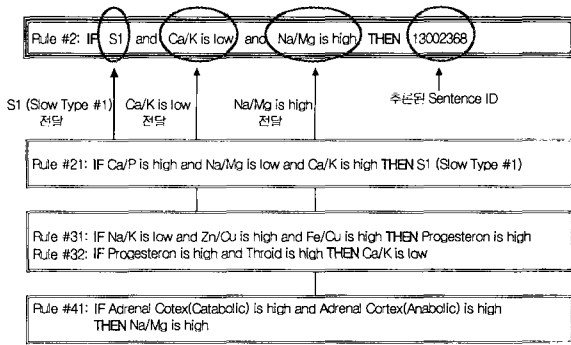


그림 9. 다중 퍼지 규칙베이스와 정보 추론 과정

따라서 환자의 다양한 미네랄 수치와 미네랄들간의 비에 따라 지능적으로 추론된 각 ID에 따른 영어문장들과 해당 문장들이 번역된 한국어 문장들이 데이터베이스로부터 추론되어 그림 8(b)와 같이 조합되어 환자와 의사에게 의료정보로써 제공된다.

4. ITS 구현

4.1 ITS 구성

본 논문에서 구현한 ITS의 구성도는 다음 그림 10과 같다. ITS에는 Analysis DB, Korean paragraph DB, Element DB, English paragraph DB 등이 있다. ITS 시스템은 미네랄 수치입력 인터페이스와 포트를 출력해주는 부분으로 크게 나누어진다. 미네랄 수치를 입력하면 입력한 수치는 수치 데이터베이스에 저장되고 그 입력된 수치값으로 문장 데이터베이스에서 문장을 추출하게 된다. ITS는 MS VC++과 MS ACCESS를 이용하여 Windows XP환경에서 구현하였다.

그림 11은 ITS 실행결과인데, 그림의 왼쪽에서 TMA자료를 입력하면 오른쪽에 의료정보자료가 한글과 영문 두가지 형태로 출력된다.

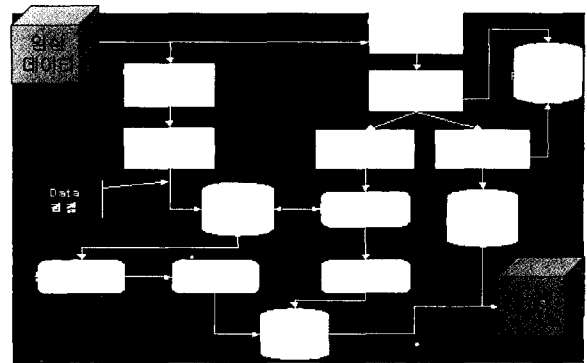
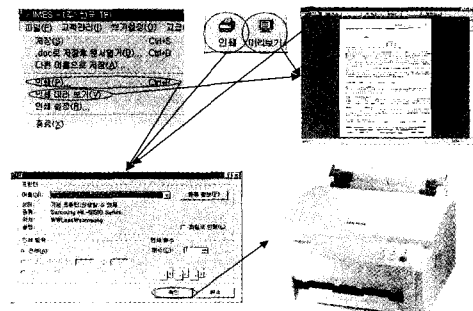
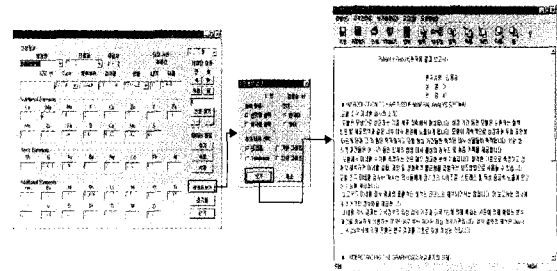


그림 10. ITS 시스템 구성도



(a) TMA 자료입력화면과 출력결과



(b) TMA 출력결과와 저장 및 출력방법

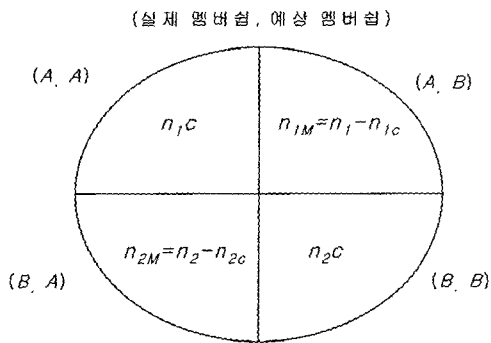
그림 11. ITS 실행결과

4.2 ITS 성능분석

분류방법의 적절성을 확인하기 위해 오분류율(misclassification probability)을 계산해야 하는데, 본 논문에서 오분류율의 계산은 표본을 이용하여 다음 식 (1)에 의해 계산하였다.

$$AER(actual\ error\ rate) = p_1 \int_{R_2} f_1(x) dx + p_2 \int_{R_1} f_2(x) dx \quad (1)$$

여기에서 R_1 과 R_2 는 표본의 크기에 의해 결정되는 분류 지역(classification regions)을 의미한다. 이것을 사용하는 이유는 위 식(1)의 방법은 미지의 밀도함수(density function), $f_1(x)$ 와 $f_2(x)$ 의 함수로 구성되어 있으므로 표본정보만을 갖고 있는 본 논문에서는 모집단의 분포에 종속되지 않고 오분류율을 계산할 수 있어야 하기 때문에 APER(APparent Error Rate)을 이용하였다.



n_{1C} = 실제 A그룹에 속하면서 A그룹으로 잘 분류된 수
 n_{1M} = 실제 A그룹에 속하면서 B그룹으로 잘못분류된 수
 n_{2C} = 실제 B그룹에 속하면서 B그룹으로 잘 분류된 수
 n_{2M} = 실제 B그룹에 속하면서 A그룹으로 잘못분류된 수
 그림 11. 오분류율 측정방법

오분류율은 $APER = \frac{n_{1M} + n_{2M}}{n_1 + n_2}$ 로 계산된다. 본 논문에서

오분류율 측정 결과 8가지 대사타입에 대해서는 오분류율이 거의 0에 가까웠으나 표본의 크기가 작은 미네랄간의 관계 등에서는 그림 12의 분포율과 같은 오분류가 발생하였는데, 표본크기가 200이상일 경우 오분류율은 거의 0에 가까운 값을 나타내었다.

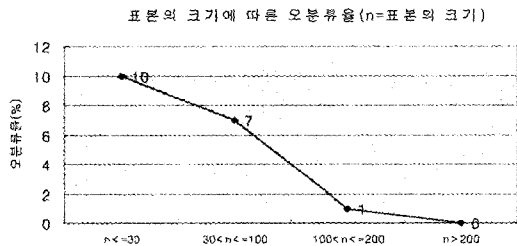


그림 12. 오분류율

다중 퍼지 규칙베이스를 구축하였을 때와 단일 퍼지 규칙베이스를 구축하였을 때의 ITS 성능분석 결과 다중 퍼지 규칙베이스에 의한 방법이 기존 단일 퍼지 규칙베이스에 의한 것보다 5배 정도 처리속도가 향상되었고, 결과의 정확도도 약 1.4배 증가하였다.

결과의 정확성은 국내 환자의 임상자료 약 2000건에 대해 측정된 것으로, 시스템에서 전체적으로 한국어로 번역되어 출력되는 최종 결과지가 미국에서 보내오는 영문으로 된 결과지와 일치하는 정도를 측정된 것이다. 환자 한 건당 결과지에 포함해서 출력해야 할 총 50개의 문장들 중 단일 퍼지 규칙 베이스로 구축한 경우는 34개 정도의 문장만이 일치하였으나 다중 퍼지규칙 베이스로 시스템을 구축한 경우 평균적으로 46개 정도 일치하였고 4개 정도 비슷하거나 잘못된 문장을 출력하였는데, 이것은 제안한 방법이 기존에 비해 약 24% 정도 정확성이 향상된 수치이다. 현재 ITS 시스템의 데이터베이스에 저장된 총 영문 종류는 각 대사 및 미네랄 타입별로 약 1000개 정도이다.

ITS의 파일럿 테스트를 위해 임상실험 결과를 중심으로 실험을 하였는데, 결과 환자로부터 모발검사를 실시한 후 미국 TEI에 의뢰하여 결과를 받은 후 환자와 담당 의사에게

돌려주는 전 과정이 ITS 사용전에는 보통 50분 정도의 시간을 소요하였는데, ITS를 사용할 경우 약 7분 정도가 소요되어 ITS를 사용하지 않은 경우보다 86% 일의능률을 가져올 수 있었다.

이상으로부터 ITS 시스템에 대한 활용도 대비 일의능률에 대한 만족도는 다음 그림 13과 같다. 이 그림에서 만족도란 활용도를 100으로 했을 때의 일의 능률 값의 비율 나타낸 것으로, '만족도 100'이란 건당 처리속도가 0임을 의미하므로 사실상 무의미한 값을 나타낸다. 이 회사에서는 환자 1건당 처리속도를 출력 속도를 감안하여 3분 정도를 가장 적당한 값으로 간주하고 있으므로, 이 논문에서 만족도는 기본값을 3분으로 하고 추가된 시간에 대해서 분석하였다. 따라서 50분, 7분에 대한 만족도를 측정하기 위해 본 논문에서 다음 식 (2)와 같은 방법을 제안하였는데, 이 식에서 50은 기존에 ITS가 없는 상태에서의 처리속도이며 3은 최소 기본 처리속도를 말하고, 3은 '만족도 100'을, 50은 '만족도 0'을 의미한다.

$$ITS \text{ 만족도} = 100 * (Max - \text{소요시간}) / (Max - Min) \quad (2)$$

여기서 Max=50, Min=3 (3 ≤ 소요시간 ≤ 50) 을 나타낸다.

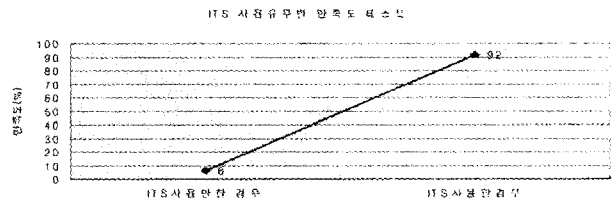


그림 13. ITS 파일럿 테스트 분석

5. 결 론

본 논문에서는 TMA를 위한 지능적 의료정보 시스템(ITS)구축하였다. 기존 TMA 분석방법의 문제점을 개선하기 위해 다단계 통계적 분류방법을 활용하여 다중으로 영향을 주는 요인들에 대하여 TMA 자료를 분류하였고, 다중 퍼지 규칙베이스를 구축하여 지능적인 퍼지추론 방법에 의해 미네랄들을 검색하고 관련성을 연구하였다. ITS의 파일럿 테스트결과, 업무능률이 ITS 시스템을 사용하지 않았을 경우보다 86%이상 향상되었고, 사용자 만족도가 92%정도 되었고, 시간당 처리하는 건수도 기존에 비해 약 7.5배이상 증가하게 되었으며 결과의 정확도도 약 92%정도로 측정되었고, 시스템의 평균 처리속도도 약 5배이상 증가하였다.

앞으로 연구해야할 과제는 좀더 많은 임상결과로부터 시스템을 개선하여 오류율을 최소화하며 웹상에서 서비스할 수 있도록 시스템을 확장하는 일이다.

참 고 문 헌

[1] Watts, D,L., Trace Elements and Other Essential Nutrients-Clinical Application of Tissue Mineral Analysis, 2nd Writer's B-L-O-C-K edition, 1997.
 [2] 의계신문(The medical World News), 1247호,

저 자 소 개

2001.1.22일자.

[3] 장수익, 김경곤, 이복기, 김형준, 유수현, 강희철, 윤방부, "당뇨환자에서의 보팔네 미네랄의 함량연구", 최신의학 45권 5호 별책, 2002.

[4] Safavian, S.R. and Landgrebe, D., "A survey of decision tree classifier methodology", *IEEE Transactions on System, Man, and Cybernetics.*, Vol.21, No.3, pp.660 674, 1991.

[5] Brown, D.E. and Corruble, V. and Pittard, C. L., "A Comparison of decision tree classifiers with backpropagation neural networks for multinomial classification problems", *Pattern Recognition*, Vol. 26, No. 6, 953 961, 1993.

[6] Johnson, R. A. and Wichern, D. W., *Applied Multivariate Statistical Analysis*, 3rd edition, Prentice Hall, 1992.

[7] Turksen, I.B. and Zhong, Z., "An Approximate Analogical Reasoning Schema based on Similarity Measures and Interval Valued Fussy Sets," *Fuzzy Sets and Systems*, North Holland, Vol. 34, pp.323 346, 1990.

[8] Lee, C.C., "Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part I", *IEEE Transactions on System, Man, and Cybernetics.*, Vol.20, No.2, pp.404 418, 1990.

[9] J.R.Graham and K.S.Decker, "Towards a Distributed Environment Centered Agent Framework", *UMASS Technical Report*, 1998.



조영임(Young Im Cho)

1988 : 고려대학교 전산학과 졸업
 1990 : 고려대학교 전산학과 석사(인공지능전공)
 1994 : 고려대학교 전산학과 박사(인공지능전공)
 1996-2005 : 평택대학교 컴퓨터학과 교수
 2005-현재 : 수원대학교 컴퓨터학과 교수
 1995-1996 : 삼성전자 멀티미디어 연구소 선임연구원
 1999-2000 : University of Massachusetts, at Amherst, Dept. of Computer Science, Post-doc
 2003-현재 : 한국퍼지 및 지능시스템학회 이사 겸 편집위원
 2003-현재 : 한국여성정보인협회 총무이사
 2003-현재 : 한국공학교육학회 편집위원
 2004-현재 : 한국전자상거래학회 학술이사
 2004-현재 : 제어자동화시스템공학회 학술이사
 2005-현재 : 한국정보과학회 편집위원

E-mail : ycho@suwon.ac.kr