TMA 분석을 위한 지능적 의학 전문가 시스템의 설계 및 구현

(Design and Implementation of an Intelligent Medical Expert System for TMA(Tissue Mineral Analysis))

조 영 임 † 한 근 식 ††
(Youngim Cho) (Geunshik Han)

요 약 모발분석(TMA: Tissue Mineral Analysis)은 머리카락 속에 있는 30여 가지의 미네랄과 8가지의 중금속의 양과 중요 미네랄 비율을 분석하여 체내에 과잉, 결핍 및 불균형 상태를 평가하고, 그 결과가 현재 인체에 미치는 영향을 예측하여, 건강을 유지하는 방향을 제시하는 임상 영양학 및 독성학 모발 조직 검사방법을 말한다.

그러나 국내 TMA 분석방법은 몇가지 문제점이 있다. 첫째, TMA 분석기기는 있으나 분석결과를 해석할 수 있는 한국형 의학 정보 데이타베이스가 없다. 둘째, 미국에서 보내오는 TMA 검사결과 자료가 영문이며 철저한 보안에 바탕을 둔 그래픽 파일 형태이므로 활용성이 적다. 셋째, TMA 관련 데이타베이스가 있어도 의료기관에서 사용하기 어려운 매우 낮은 수준이므로 TMA 분석 및 의료서비스를 위해 매번 미국에 의뢰해야 하므로 심각한 외화낭비를 초래한다. 넷째, TMA 결과가 서구식 생활패턴에서 비롯된 데이타베이스로부터 구축된 것이므로 검사결과의 신뢰성 문제가 발생한다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해 국내 전문 기관으로부터 자료를 제공받아 TMA 관련 국내 최초 지능적 의학 전문가 시스템(IMES: Intelligent Medical Expert System)을 개발하였다. IMES는 TMA 자료를 다단계 통계분석 방법에 의한 결정트리 분류기를 이용하여 분류하고 다중 퍼지 규칙베이스를 구축하여, 지능적 퍼지추론 방법에 의해 한글화된 데이타베이스로부터 복잡한 자료를 추론하도록 구축하였다. 본 IMES 시스템을 실제적용한 결과 업무능률과 만족도가 각각 86%, 92% 증가함을 알 수 있었다.

키워드 : 지능적 의학 전문가 시스템, 모발 분석(TMA: Tissue Mineral Analysis), 결정트리 분류기, 퍼지 규칙베이스, 퍼지추론

Abstract Assesment of 30 nutritional minerals and 8 toxic elements in hair are very important not only for determining adequacy, deficiencies and unbalance, but also for assessing their relative relationships in the body. A test has been developed that serves this purpose exceedingly well. This test is known as tissue mineral analysis(TMA). TMA is very popular method in hair mineral analysis for health care professionals in over 46 countries' medical center.

However, there are some problems. First, they do not have database which is suitable for korean to do analyze. Second, as the TMA results from TEI-USA is composed of english documents and graphic files prohibited to open, its usability is very low. Third, some of them has low level database which is related to TMA, so hairs are sent to TEI-USA for analyzing and medical services. it bring about an severe outflow of dollars. Finally, TMA results are based on the database of american health and mineral standards, it is possibly mislead korean mineral standards.

The purposes of this research is to develope the first Intelligent Medical Expert System(IMES) of TMA, in Korea, which makes clear the problems mentioned earlier. IMES can analyze the tissue mineral 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.

† 종신회원 : 평택대학교 컴퓨터과학과 교수
yicho@ptuniv.ac.kr
†† 비 회 원 : 한신대학교 정보시스템 공학과 교수
gshan@hanshin.ac.kr
논문접수 : 2002년 8월 2일
심사완료 : 2003년 10월 30일

Pilot test of this systems are increased of business efficiency and business satisfaction 86% and 92% respectively.

Key words : Intelligent Medical Expert System, Tissue Mineral Analysis, Decision Tree Classifier, Fuzzy Rulebase, Fuzzy Inference

1. 서 론

최근 자료에 의하면 사망원인이 1990년대에 비해 만성 감염성 질환(폐렴, 결핵 등)에서 만성 대사성 질환(심장질환, 암 등)으로 변화하고 있다고 한다. 이러한 사망순위에 대한 원인은 탄수화물, 지방, 단백질 등 칼로리 영양소가 비타민, 미네랄, 섬유질 등 대사 영양소보다 과다 섭취되는 경향으로 인해 영양 불균형을 초래하기 때문이라고 한다[1,2].

인체의 영양상태를 스크리닝하는 방법은 크게 3가지가 있는데 혈액검사, 요 검사, 모발분석(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 모발분석에 의한 검사는 세포내의 대사상태를 가장 정확하게 반영하고 있어서 세계 10대 사망원인을 정확하게 측정하고 있으며 이미 세계 48개국에서 도입되어 사용되는 보편적인 방법이다.

TMA를 하는 국외 대표적인 기업으로는 미국 TEI (Trace Elements Inc.)가 있다. 이 회사는 약 20여년간 임상실험한 약 35만건 이상의 자료로부터 구축된 데이타베이스를 보유하고 있고, 현재 미네랄 관련 가장 많은 임상자료를 보유하고 있으며 세계 40여개국에 있는 의료기관에 의학 서비스를 제공하는 FDA 승인된 세계 최대의 TMA 전문기관이다[3].

현재 국내에서는 TMA를 하는 기관으로는 (주)한국티이아이[4]와 (주)메디넥스[5] 등 2곳이 있다. (주)한국티이아이는 미국 TEI에 결과를 의뢰하여 분석된 자료를 국내 의료기관등에 제공하는 유일한 기관이나 자체 데이타베이스는 보유하고 있지 않다. (주)메디넥스는 미국 TEI에 의뢰하지 않고, 중금속 오염도 측정 등 집단 검사를 주로하고 있으나 검사수준이 TEI보다 훨씬 떨어진다. 또한 자체 데이타베이스를 보유하고 있으나 분석 수준이 매우 낮아 의료기관에서 사용하기에는 적절하지 못하다.

국내에서 TMA 임상실험을 하는 의료기관으로는 고려대학교병원, 연세대학교 세브란스 병원, 분당 차병원, 삼성제일병원, 경희대 병원, 카톨릭 대학병원 등을 중심으로 미네랄에 대한 인식이 확산되고 있고 국민의료보험 적용여부에 관해서도 논의 중이므로 TMA 방법의 활성화가 조기에 달성될 수 있을 것으로 전망한다.

그러나 현재 국내 TMA 분석방법은 분석결과를 해석할 수 있는 의학 전문 프로그램이 개발되어 있지 않고, 일부 국내기업에서 기구축한 TMA 관련 데이타베이스가 매우 낮은 수준이어서 의료기관에서 충분한 의료서비스를 제공하기 어렵다. 그러므로 현재 국내에서는 대부분의 TMA 분석을 미국에 의뢰하게 되는데 이 결과가 서구식 생활패턴에서 비롯된 데이타베이스로부터 구축된 것이므로 한국인에 대해서는 보다 분석적인 한국형 데이타베이스 및 TMA 결과분석이 필요하다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해 TMA 분석을 위한 알고리즘을 개발하고 TMA 분석결과에 대한 데이타베이스를 구축하여 지능적 의학 전문가 시스템(IMES: Intelligent Medical Expert System)을 구축하고자 한다. 이것을 위해 국내 TMA 전문기관인 (주)한국 티이아이에서 미네랄 관련 임상자료를 이용하였으며 통계분석적 오류와 데이타베이스 구축 및 추론상의 오류를 최소화하기 위해 다중 퍼지규칙 베이스를 구축하였고 다단계 통계분류 방법을 이용하였다. 또한 한글화 데이타베이스 구축을 위해 한글화 작업도 병행하였다.

본 논문의 구성은 다음과 같다. 2장에서는 모발분석의 정의, 필요성 및 의학적 배경 등을 설명하고, 3장에서는 본 논문에서 제안하는 지능적 추론방법, 다중 퍼지규칙 베이스 구성 및 통계적 분류 방법을 설명하며, 4장에서는 IMES 시스템의 개요와 주요 기능 설명 및 효율성 등을 분석하고, 5장에서 결론을 내리고자 한다.

2. TMA(모발분석)

2.1 정의 및 필요성

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[1]. 또한 호르몬의 세포수준 활성을 제시(갑상선 및 부신)하며 영양, 신경 및 호르몬의 불균형을 통해 질환발생 가능성을 설명한다.

일반적으로 모발조직 검사방법은 조직 미네랄 검사와 모발 미네랄 검사의 두 가지로 분류된다. 첫째, 조직 미네랄 검사(TMA: Tissue Mineral Analysis)는 조직(머리카락, 일반 조직) 중 미네랄의 결핍, 과잉 및 불균형을 평가하는 방법을 말하며 머리카락만을 갖고 하는 것보다 좀더 포괄적인 의미를 갖는다. 따라서 모발분석이라 하면 일반적으로 조직 미네랄 검사를 의미한다. 둘째, 머리카락 미네랄 검사(HTMA: Hair Mineral Analysis)는 비용 대비 효율이 가장 높고 정확한 TMA가 HTMA인데, 일반적으로 TMA는 머리카락을 이용한 조직 미네랄 검사[6,7]를 많이 하고 있으므로 HTMA를 TMA라고 하기도 한다.

2.2 특징 및 장단점

TMA의 특징 및 장단점은 다음과 같다. 첫째, 모발 표본은 비침해 방법으로 채취되므로 혈액, 소변, 기타 다른 조직보다 빠르고 쉽게 수집할 수 있다. 둘째, TMA 방법은 다른 방법의 미네랄 분석보다 비용 대비 효율이 우수하다. 셋째, 모발은 혈액과는 달리 미량원소의 농도에 민감한 영향을 미치는 항상성 기전에 덜 민감하다. 넷째, 미네랄의 저류 또는 손실과 같은 장기간의 편차들은 혈액보다 모발조직에서 쉽게 검출된다. 다섯째, 대부분의 모발조직 미네랄 농도는 혈액이나 다른 조직에서 발견되는 미네랄보다 농도가 매우 높다. 여섯째, 모발조직은 과거뿐만 아니라 현재의 미량원소의 수치기록을 제공한다. 일곱째, 혈청뿐 아니라 외부원인으로부터 모발로 들어가는 물질의 정보를 제공한다. 여덟째, 독성 중금속 수치를 평가하는데 유용하다[8].

2.3 혈액측정방법과의 차이점

혈액이나 혈장에는 각종 미네랄이 함유되어 있으나 이들이 인체 미네랄 상태를 완전하게 표시할 수 없다고 한다[9]. 그 이유는 미네랄은 항상성의 일환으로 미네랄의 혈장 수치는 조직 미네랄에 의해 유지되고 있으나, 혈장농도는 사람의 기분이나 혈액채취시간 또는 표본채취적의 음식 섭취에 따라 변동이 다양하기 때문이다. 일반적으로 혈액은 인체의 저장 부위로부터 또는 저장부위로의 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. IMES: 지능적 의학 전문가 시스템

3.1 연구의 필요성

일반적인 TMA 분석과정은 다음과 같다. 먼저, 환자의 모발 약 5그램을 채취하여 TMA 분석기기에 넣어 환자 개개인에 대해 38종 이상의 미네랄 분포량을 얻는다. 이로부터 환자의 현재 영양상태 및 인체의 길항작용 등을 설명하게 되며, 예방 및 치료할 수 있는 다량의 영양자료와 식단을 데이타베이스로부터 분석하여 20여페이지에 달하는 결과지 형식을 갖추어서 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.

그러나 국내에서는 TMA 자체분석 방법이 없기 때문에 미국 TEI에 의뢰하고 있는데 검사하려는 자료를 미국 TEI에 보내면, 자체 대규모 데이타베이스로부터 개개인의 미네랄 양과 비, 대사 및 영양상태 등을 종합적으로 분석한 영문으로 된 환자 및 의사용 검사결과지를 국내 의뢰기관에 제공하게 된다.

TMA에 관해서는 카톨릭 대학병원, 분당 차병원, 고려대학교 병원, 삼성제일병원, 연세대학교 세브란스 병원[10] 피부과 및 50여개 종합병원과 지방 350여 병원에서 전문의료진들이 TMA 및 임상영양에 대한 필요성을 인식하여 임상 중이므로 데이타베이스 구축에 관한 연구는 필요하다.

현재 국내 TMA 분석 방법에서의 문제점은 다음과 같다.

첫째, TMA 분석 기계는 있으나 분석결과를 해석하여 환자와 의사들에게 제공할 수 있는 프로그램이 없다. 현재는 TMA 분석을 위해 미국 본사에 의뢰하여 결과를 받아보는 방법을 이용하고 있는데, 이 방법은 현재 국내에서 TMA 분석을 의뢰하는 수요자가 증가하는 추세를 감안할 때 시간과 비용 측면에서 적절하지 못하다.

둘째, 한국에 미국과 같은 의학정보 데이타베이스가 없다. 따라서 TMA 분석기기가 문제가 아니라 분석방법(프로그램, 데이타베이스) 개발이 필요하다. 이 의학정보 데이타베이스는 미국 본사만의 노하우이므로 제공되지 않는다.

셋째, 미국에서 보내오는 TMA 검사결과가 영문이라 가독성이 떨어지므로 한글화된 결과지가 필요하며 철저한 보안을 바탕으로 한 그래픽 파일형태이므로 일회성이라 재활용이 불가능하다.

넷째, 검사비용으로 인한 외화낭비가 초래된다. 즉, 검사 한건당 15만원인데 이중 미국 TEI 본사에 5만원을 지급해야 한다고 하므로 현재 년간 18억원 이상의 심각한 외화낭비가 발생한다.

다섯째, 검사결과의 신뢰성이 부족하다. 이것은 한국적 특성에 맞는 데이타가 아니라 미국의 데이타베이스에 의한 결과물이므로 검사결과의 신뢰성이 부족할 수 밖에 없고 한국적 임상 결과에 대한 검증자료가 없다.

따라서 본 논문에서는 미국에 의뢰하지 않고 국내에

있는 TMA 분석기기로도 충분히 미국에서와 같이 분석 및 해석이 가능할 뿐 아니라 한국어로 제공이 가능한 한국형 데이타베이스 시스템(지능적 의학 전문가 시스템: IMES)을 구축하고자 한다. 이렇게 함으로써 미국에 검사비용 및 결과지 해석비용으로 지불하던 외화 낭비를 줄이고 환율 및 부대비용의 지불로 인하여 환자의 의료비를 절감하도록 하여 미국 TEI의 독점을 방지하고자 한다. 이러한 TMA 관련 데이타베이스 구축에 관한 연구는 국내에서는 최초로 시도되는 연구이다.

3.2 연구내용

3.2.1 연구개요

본 논문에서 연구하려는 TMA 분석을 위한 지능적 의학 전문가 시스템의 연구목표는 그림 1과 같이 TMA 자료의 통계적 분석 및 그룹화와 사용자 중심의 정보검색이 가능한 데이타베이스 구축에 관한 것이다.

본 논문에서 제안하는 연구내용 및 방법은 다음 그림 2와 같다. 통계분석단계에서 통계적 분석이 완료된 후 시스템 구축단계에서 TMA 데이타베이스를 구축하고,

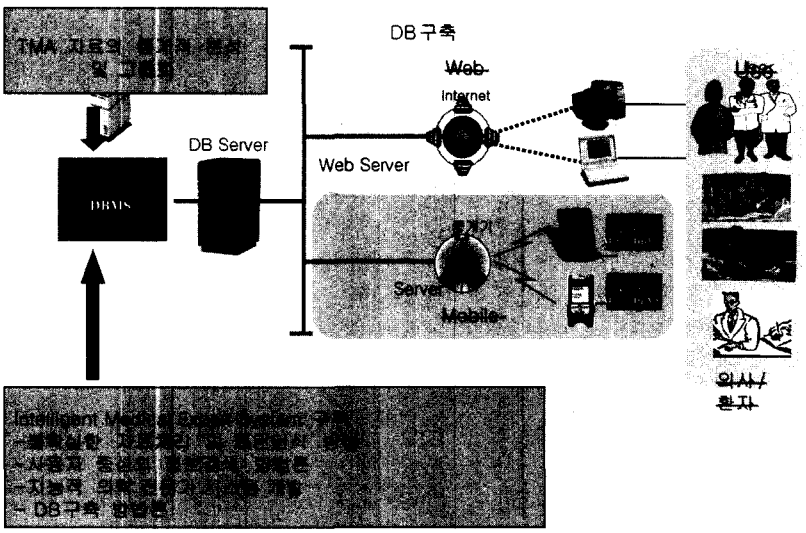


그림 1 연구목표 및 주요내용

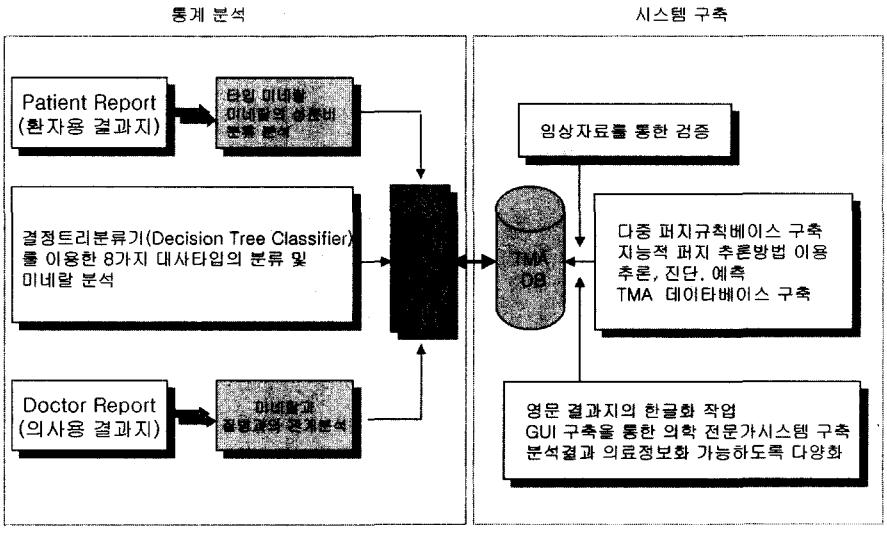


그림 2 연구방법

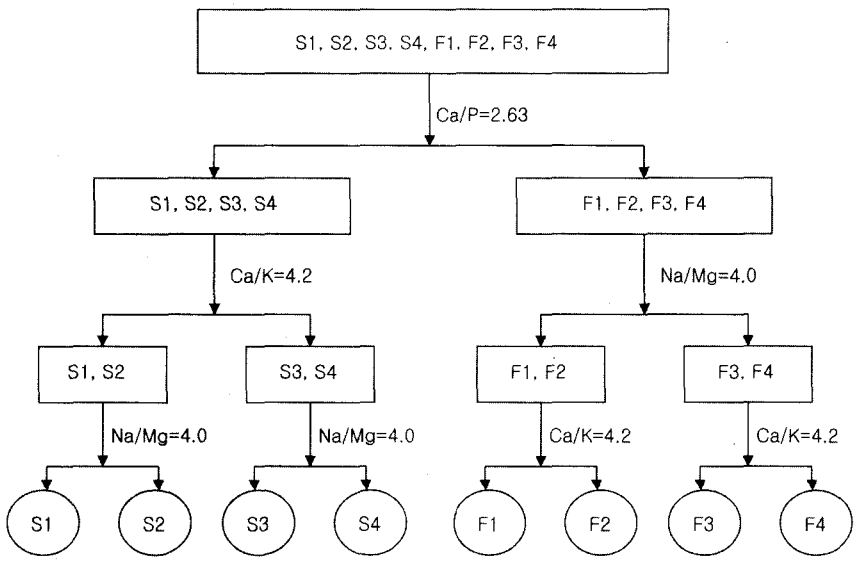


그림 3 결정트리에 의한 TMA 타입 분류

임상자료를 통해 시스템을 검증한다.

3.2.2 통계적 분류 및 분석 방법

지능적 추론의 배경이 되는 대사타입의 분류 및 미네랄의 분석은 통계적 분류 및 분석 방법 중에서 결정트리분류기(Decision Tree Classifier)를 이용하여 TMA 자료의 대분류에 해당하는 인체의 8가지 대사타입(빠른 대사 4타입, 느린 대사 4타입)을 분류하고자 한다. 이 방법을 사용하는 이유는, 1단계 분류기는 클래스간의 변별력 중에서 평균이 최대가 되는 특징조합을 이용하나, 이 특징조합이 모든 클래스에 대하여 최적의 변별력을 항상 갖는 것이 아니므로 두 종류이상의 결정을 적절한 수순에 따라 몇 번이고 반복하여 입력패턴의 클래스를 결정해야 하므로 본 논문에서는 다단계분류기인 결정트리 분류방법[11-13]을 응용하여 분류하고자 한다.

위 그림 3은 결정트리분류기가 패턴집합을 분류한 것인데, 그 과정은 다음과 같다. 우선 루트노드에서 미네랄 비 Ca/P(=2.63)를 이용하여 느린 대사타입(Slow Metabolic Type)의 패턴집합을 $\{S_1,S_2,S_3,S_4\}$와 빠른 대사타입(Fast Metabolic Type) 패턴집합 $\{F_1,F_2,F_3,F_4\}$의 서브패턴집합으로 분류한 다음, 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 분류하였다. 마지막 단계에서 Na/Mg(=4.0)를 이용하여 2차 분류된 패턴집합$\{S_1,S_2\}$와 $\{S_3,S_4\}$를 S_1과 S_2 그리고 S_3과 S_4로 최종 분류하고, 같은 방법으로 Ca/K(=4.2)를 이용하여 2차 분류된 패턴집합 $\{F_1,F_2\}$와 $\{F_3,F_4\}$를 F_1과 F_2 그리고 F_3와 F_4로 최종 분류하여 결정트리를 구성하였다.

대사타입의 분류는 특성변수에 대한 클래스별 분포를 이용해도 같은 결과를 보여 준다. 다음은 특성변수를 계급적으로 적용하여 클래스별 분포를 이용한 대사타입의 분류를 보여준다.

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 4이다.

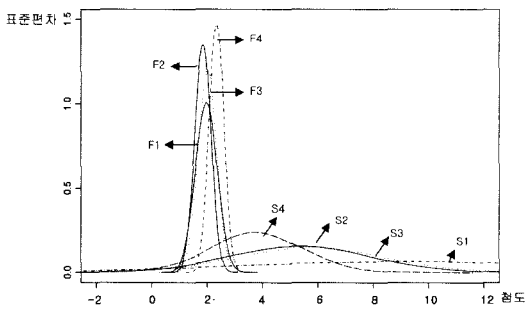


그림 4 Ca/P 클래스 분포를 이용한 대사타입의 분류

위 그림에서 왼쪽의 4개 분포는 첨도가 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 첨도가 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.

위와 같이 분류된 대사타입 S를 Ca/K 클래스 분포에 의해 분류하면 다음 그림 5와 같다.

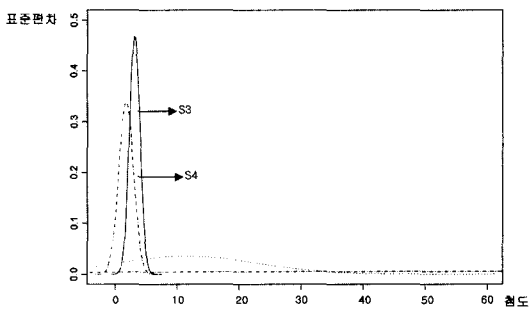


그림 5 대사타입 S의 Ca/K 클래스 분포에 의한 분류

위 그림을 살펴보면 Ca/K의 분포는 첨도가 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 첨도가 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.

이제 Ca/K의 분포에 의해 분류된 S_3와 S_4 그리고 S_1과 S_2를 Na/Mg의 분포에 의해 분류하면 각각 그림 6, 그림 7과 같다. 그림 6에서 Na/Mg분포를 보면 왼쪽의 높은 첨도와 낮은 표준편차를 갖는 분포는 $S_3 \sim (\hat{\mu}=2.7, \hat{\sigma}=0.76)$이고 아래쪽의 분포는 표분편차가

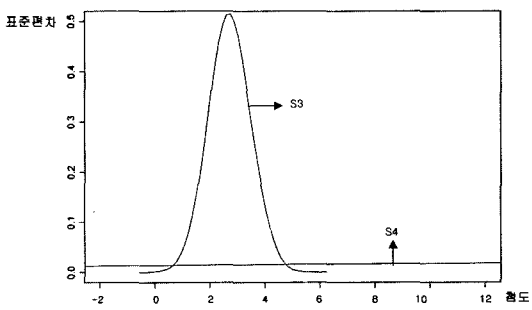


그림 6 S_3, S_4의 Na/Mg 클래스 분포에 의한 분류

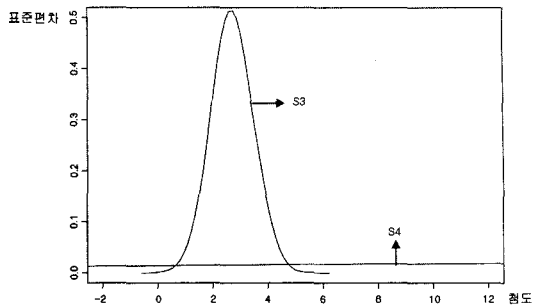


그림 7 S_1, S_2의 Na/Mg 클래스 분포에 의한 분류

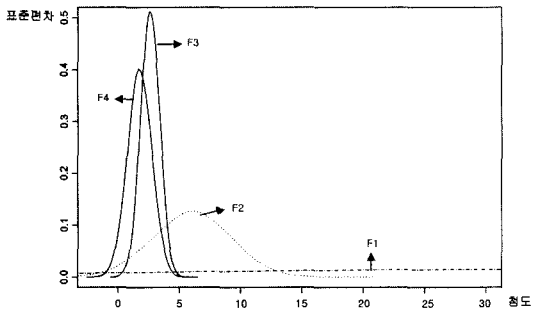


그림 8 대사타입 F의 Na/Mg 클래스 분포에 의한 분류

매우 큰 $S_4 \sim (\hat{\mu}=15.3, \hat{\sigma}=21.27)$로 분류됨을 알 수 있다. 마찬가지로 그림 7에서 Na/Mg분포를 보면 왼쪽의 높은 첨도와 낮은 표준편차를 갖는 분포는 $S_2 \sim (\hat{\mu}=1.25, \hat{\sigma}=0.97)$이고 오른쪽의 분포는 표분편차가 크고 첨도도 낮은 $S_1 \sim (\hat{\mu}=6.16, \hat{\sigma}=3.33)$으로 분류됨을 알 수 있다.

대사타입 F의 Na/Mg 클래스 분포에 의한 분류는 그림 8과 같다.

위 그림 8에서 왼쪽의 첨도가 높은 분포가 $F_4 \sim (\hat{\mu}=1.78, \hat{\sigma}=0.95)$와 $F_3 \sim (\hat{\mu}=2.62, \hat{\sigma}=0.75)$이고 표준편차가 큰 아래쪽의 분포는 $F_2 \sim (\hat{\mu}=6.04, \hat{\sigma}=3.07)$와 $F_1 \sim (\hat{\mu}=29.6, \hat{\sigma}=27.33)$이다. 끝으로 첨도가 높은 F_3와 F_4는 Ca/K 클래스에 의한 분포로 그림 9와 같이 분류되고 F_1과 F_2는 Ca/K 클래스에 의해 그림 10과 같이 분류된다.

TMA에서 가장 기본이 되는 대사타입의 분류 후에 각 대사타입 별 영양미네랄과 독성미네랄의 영양상태와의 관계는 상관분석을 통해 상관계수가 0.3보다 큰 미네랄들을 대상으로 분석하였다.

3.2.3 다중 퍼지 규칙베이스 설계

분석된 통계를 바탕으로 본 논문에서는 불확실한 다

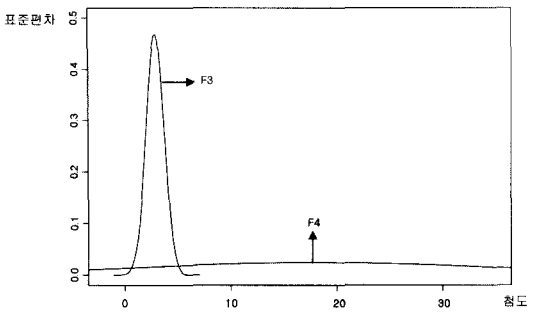


그림 9 Ca/K 클래스 분포에 의한 F_3, F_4의 분류

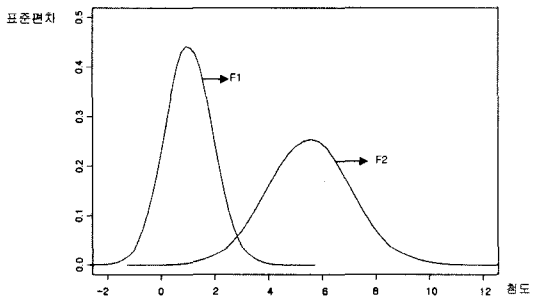


그림 10 Ca/K 클래스 분포에 의한 F_1, F_2의 분류

량의 정보로부터 지능적 추론이 가능한 데이타베이스 구축을 위해 다중 퍼지 규칙베이스 방식을 제안한다. 다중 퍼지 규칙베이스는 TMA 자료와 같이 다양한 미네랄에 의해 의사결정을 해야하는 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[14] 자료처리 및 제어시스템에 적당하기 때문이다.

본 논문에서는 다중 퍼지 규칙베이스 구축을 위해 다음과 같은 두 단계를 갖는데, 제1단계 구축 후 제2단계를 통해 다중 퍼지 규칙베이스 구축이 완성된다.

[제1단계] 대사타입 분류에 의한 퍼지 데이타베이스 구축

제1단계는 빠른 대사와 느린 대사를 결정짓는 미네랄들의 상관관계 분석으로부터 대사타입에 따른 퍼지 데이타베이스를 구축하는 과정을 말한다. 이것을 위해 본 논문에서는 영양미네랄(Nutritional Mineral) 30종과 독성 미네랄(Toxic Mineral) 8종류 등 38종 미네랄과 영양상태와의 관계, 영양미네랄들의 비(ratio), 독성미네랄들의 비(ratio)와 영양상태와의 관계, 그리고 TMA를 통한 각종질환과 미네랄과의 경향에 관해 분석한 결과 다음 표 1을 얻었다.

표 1에서 문장 분석시 영향을 미치는 주요 미네랄 비율을 나타내는 언어적 변수인 "높다(high)" "낮다(low)" "적당(acceptable)" 등은 다음 그림 11과 같이 퍼지 집

표 1 대사타입에 따른 퍼지 규칙베이스 구성(제1단계)

분류기준	Adrenal Function			Thyroid Function		대사타입
Ca/P>2.63	부교감신경발달	Na/Mg=4.0	값이 낮으면 기능저하	Ca/K=4.2	값이 높으면 기능저하	느린 대사타입 #1
	부교감신경발달	Na/Mg=4.0	값이 높으면 기능항진	Ca/K=4.2	값이 높으면 기능저하	느린 대사타입 #2
	부교감신경발달	Na/Mg=4.0	값이 낮으면 기능저하	Ca/K=4.2	값이 낮으면 기능항진	느린 대사타입 #3
	부교감신경발달	Na/Mg=4.0	값이 높으면 기능항진	Ca/K=4.2	값이 낮으면 기능항진	느린 대사타입 #4
Ca/P=2.63	적당 (기준값)					
Ca/P<2.63	교감신경 발달	Na/Mg=4.0	값이 높으면 기능항진	Ca/K=4.2	값이 낮으면 기능항진	빠른 대사타입 #1
	교감신경 발달	Na/Mg=4.0	값이 높으면 기능항진	Ca/K=4.2	값이 높으면 기능저하	빠른 대사타입 #2
	교감신경 발달	Na/Mg=4.0	값이 낮으면 기능저하	Ca/K=4.2	값이 낮으면 기능항진	빠른 대사타입 #3
	교감신경 발달	Na/Mg=4.0	값이 낮으면 기능저하	Ca/K=4.2	값이 높으면 기능저하	빠른 대사타입 #4

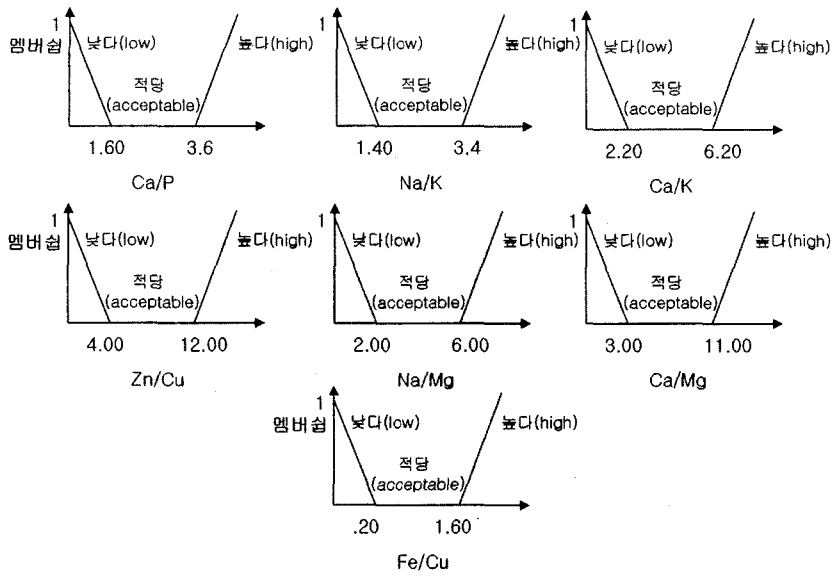


그림 11 미네랄 비율의 퍼지 집합 표현

합[15]으로 표현하였고, x축은 미네랄 비율이며 y축은 퍼지 멤버쉽 값을 나타내는데 이러한 수치들은 3.2.2절의 통계적 분석에 의해 획득하였다.

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[16].

[제2단계] 각 미네랄 양과 비율에 따른 해당 문장 데이타베이스 검색 및 출력

제2단계는 제1단계의 대분류 과정을 통해 구축한 대사타입에 따른 퍼지 데이타베이스를 세분화하고 또한 내분비선과 미네랄과의 상관관계에 대한 데이타베이스를 구축하는 과정을 말한다. 제2단계 분류 및 퍼지 규칙 베이스 설계를 위해 다음 표 2와 같은 분석 결과를 이용하였다. 예를 들면, Ca/P, Ca/Mg, Ca/Na, Ca/K, Ca/Fe가 증가하면 Parathyroid가 증가하게 되고, 반대로 Fe/Cu가 증가하면 Parathyroid가 감소함을 의미한다.

다음 표 3은 미국에서 보내주는 환자의 대사타입에 따른 결과지에 포함되어 상태를 설명하는 각 영어 문장(paragraph)과 특정 미네랄 분포와 상관관계가 있다는 사실로부터 문장을 분석하는 과정의 일부를 나타낸 것이다. 여기서 왼쪽 수치값은 분석을 위해 본 논문에서 임의로 부여한 각 영어문장 ID이다. ID는 총 8자리로 구성되며 앞의 두자리는 문장 번호들의 대분류와 관련된 문장들의 번호이며, 다음 3자리는 미네랄의 순서 예를 들어, 'Na', 'Mg', 'Ca' 순으로 미네랄에 대한 설명이 나오면 'Na'는 '001'이 되고, 'Ca'는 '003'이 된다. 또한 다른 3자리는 미네랄을 설명하는 문장들의 종류번호이다. 예를 들어, 공통문장에서 'Ca'을 설명하는 15번째 문장의 ID는 '13003015'가 된다. 이 표에서 본 바와 같이

표 2 미네랄 비율과 내분비선과의 증감관계(제2단계)

Parathyroid	Ca/P Ca/Mg Ca/Na Ca/K Ca/Fe	Fe/Cu
Thyroid	Na/Mg Fe/Cu	Ca/P Ca/K
Adrenal Cortex(Anabolic)	Na/K Na/Mg Fe/Cu Ca/Mg	Ca/P Ca/K Ca/Na
Adrenal Cortex(Catabolic)	Fe/Cu Na/Mg	Na/K Ca/K Ca/P Ca/Mg Ca/Na
Pancreas	Ca/P Ca/Mg Ca/K Ca/Fe Ca/Na	Zn/Mg Fe/Cu
Estrogen	Ca/Mg Ca/P Ca/K Ca/Na Ca/Fe Na/K	Zn/Cu Fe/Cu
Progesteron	Zn/Cu Fe/Cu	Na/K Ca/K

표 3 미네랄의 분포에 의한 문장 분석 과정의 예

문장 ID	미네랄 분포
13009014	Fe <= 0.4 And Mg < 3 And CaP < 2.63 And NaMg >= 4 And CaK < 4.2
13010001	0.2 < Mn < 4.2
13010002	0.16 < Mn <= 0.19
13010003	0.2 < Mn < 4.2
13010004	0.006 <= Mn <= 0.01
13011001	0.1 <= Cr <= 0.13
13011002	0.14 <= Cr
13011003	Cr <= 0.02 And Mn < 0.07 And 0.03 < Se < 0.11 And CaP > 2.63 And CaK >= 4.2 And NaMg < 4
13012001	0.4 <= Se
13012002	Se < 0.03 And Hg < 0.09
13012003	0.22 <= Se <= 0.31
13012004	Se < 0.03 And Hg < 0.09
13013001	1.36 <= B <= 2.0
13013002	2.0 < B
13013003	2.0 < B
13013004	2.0 < B
13014001	0.01 <= Co <= 0.017 And CaP >= 2.63 And CaK >= 4.2 And NaMg < 4
13014002	0.004 <= Co <= 0.007
13014003	0.05 < Co And CaP >= 2.63 And CaK >= 4.2 And NaMg < 4
13014004	0.01 <= Co And CaP >= 2.63 And NaMg < 4
13014005	0.05 < Co And CaP >= 2.63 And CaK >= 4.2 And NaMg < 4
13014006	0.05 < Co And CaP >= 2.63 And CaK >= 4.2 And NaMg < 4
13014007	0.013 <= Co And CaP >= 2.63 And NaMg < 4
13015001	0.015 <= Mo < 0.025
13015002	Mo = 0.001 And Age > 12
13015003	0.011 <= Mo < 0.015
13015004	0.016 <= Mo <= 0.054 And Age < 12
13015005	0.011 <= Mo < 0.015
13015006	0.002 <= Mo <= 0.009 And 12 < Age And 3000 <= CuMo < 4000 And Hg < 0.09
13015007	Mo <= 0.001 And Age < 12
13016001	S < 3000
13017001	Ge <= 0.005
13017002	0.03 <= Ge And 47 <= Age
13017003	0.03 <= Ge And 47 <= Age
13017004	0.003 <= Ge < 0.006 And Age < 12
13017005	no data
13017006	no data
13017007	Ge > 0.014 And Age >12 And CaP > 2.63 And CaK < 4.2 And NaMg >= 4
13018001	0.43 <= Ba
13018002	no data
13019001	0.059 < Bi
13020001	0.032 < Rb

영어 각 문장은 특정 미네랄의 양, 다른 미네랄과의 관계, 주요 미네랄의 비 등과 매우 밀접한 관계를 갖고 있고, 동일 미네랄 그룹내 ID 문장간 연관성은 전문가의 의견을 바탕으로 연관성을 고려하여 설계하였다. 이러한 2단계 과정을 모두 거친 후 다중 퍼지 규칙베이스 구축이 완성된다.

다음은 다중 퍼지 규칙베이스에 의한 규칙 형태와 기존의 단순 퍼지 규칙베이스에 의한 규칙형태를 비교한 예이다. 이 예에서 보는 바와 같이 본 논문에서 적용하는 다중 방법은 퍼지 규칙의 조건부에 조건명제들의 수를 3개 이하로 제한시켜 데이타베이스에서의 추론 방법을 효율화했으며 계층적 구조를 이용하여 연관된 다음 규칙과의 관계로부터 규칙의 조건부가 추론되도록 하였다.

[다중 퍼지 규칙베이스에서의 규칙형태의 예]

if Ca/P가 높고(high) Na/Mg가 낮고(low) Ca/K가 높으면(high)

then Slow type #1이다.

if Na/K가 낮고(low) Zn/Cu가 높고(high) Fe/Cu가 높으면(high)

then Progesteron을 증가시킨다.

if Progesteron이 높고(high) Thyroid가 높으면(high) then Ca/K는 낮다(low).

[단순 퍼지 규칙베이스에서의 규칙형태의 예]

if Ca/P가 높고(high) Na/Mg가 낮고(low) Ca/K가 높고(high) Na/K가 낮고(low) Zn/Cu가 높고(high) Fe/Cu가 높고 Progesteron이 높고(high) Thyroid가

높으면 (high)

then Ca/K는 낮다(low).

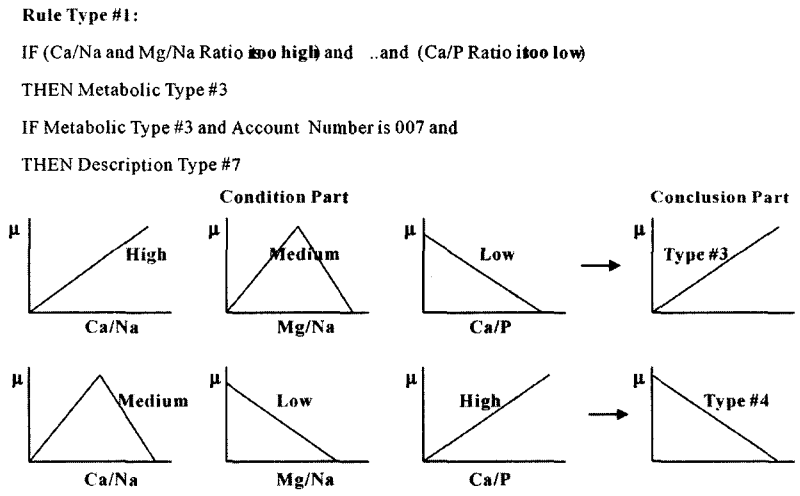
여기서 데이타베이스에는 Ca/Na, Na/Mg 값 등이 그림 11과 같이 수치적 자료로써 테이블 형태로 저장되어 있고, 높고 낮음 등 언어적 변수 등을 나타내는 함수들이 저장되어 있어서 자연어 처리 표현에 효과적으로 이용할 수 있다. 이러한 함수들은 내장된 함수를 이용하거나 수학적 함수에 의해 프로그래밍 하였다.

3.2.4 지능적 추론 방법

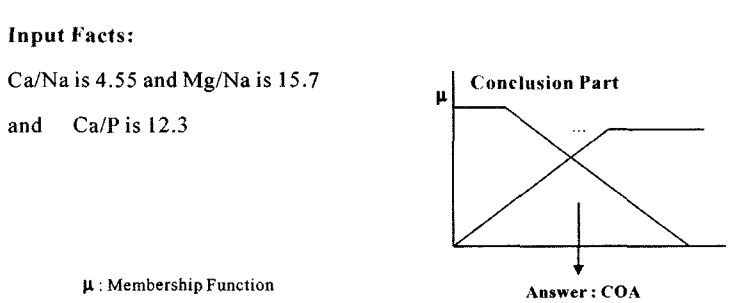
본 논문에서 구축한 지능적 퍼지 추론 방법을 위한 퍼지 규칙베이스 형태 및 추론방법은 다음 그림 12와 같다. 본 논문에서는 규칙의 조건부에는 미네랄의 비를 이용하여 구축하였고, 결론부에는 대사타입 및 규칙의 조건부에 입력되는 미네랄의 비를 이용하였다. 본 논문에서 현재까지 구축한 규칙은 약 1000개 정도이다.

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)[17]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. TAEMS에는 분산 목표들간의 스케쥴링, 목표트리 분해 및 분석 방법 등을 포함하는 새로운 개념으로 미국의 U.of Massachusetts 멀티에이전트 연구실에서 연구되었다.

다음 그림 13은 다중 퍼지 규칙베이스에서 규칙들간의 상호의존성을 TAEMS 방식을 적용해서 나타낸 그림이며, 본 논문에서 구현한 퍼지 규칙베이스의 구성형



(a) 퍼지 규칙베이스 설계



(b) 퍼지추론 방법

그림 12 퍼지 규칙베이스에서의 퍼지 추론 방법

<Fuzzy Rule Base 구성 형태>

Rule 1: IF x1 is A1 and y1 is B1 and ....z1 is C1 **THEN** k1 is R1
Rule 2: IF x2 is A2 and y2 is B2 and ....z2 is C2 **THEN** k2 is R2
.....
Rule n: IF xn is An and yn is Bn and ....zn is Cn **THEN** kn is Rn

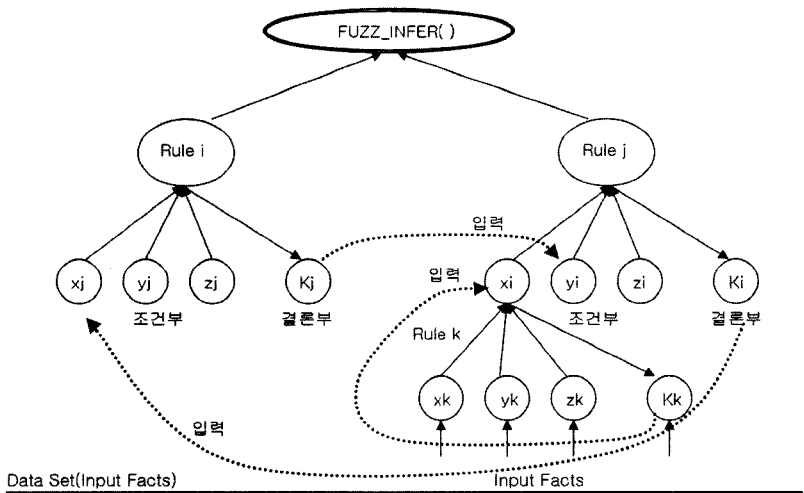
여기서 x1,y1,z1,k1 : 미네랄, 미네랄 비, 내분비선 등
A1,B1,C1,R1: 퍼지집합(high,low,medium...)

<FUZZ_INFER( ) 알고리즘>

Enter of input facts to Rule i (초기입력) ;
TAEMS에 의한 퍼지 규칙 연관관계 트리구조 생성 ;
Rule i 와 연관된 규칙들의 집합 생성; // 규칙들간 Complexity, Scheduling은 TAEMS 방식 따름
퍼지추론 할 수 있는 입력형태로 변환;

If input fact == conditional part of Rule m(연관된 규칙)
then FUZZ_INFER_I(RuleBase, input facts);
// 입력이 규칙의 조건부와 모두 일치하고, 결론부가 단일일 경우
즉, 규칙들간의 다중 상호 의존관계가 없는 경우 퍼지추론하는 과정

If input fact == some conditional part of Rule n(연관된 규칙)
then FUZZ_INFER_II(RuleBase, input facts);
{ for(i=0; i<=n; i++) // m, n은 TAEMS에 구축된 연관 규칙들의 갯수
for(j=0; j<=m; j++)
{ x[i] of Rule[i] = K[j] of Rule[j];
y[i] of Rule[i] = K[k] of Rule[k];
x[i] of Rule[i] = K[m] of Rule[m];
}
// 입력이 규칙의 조건부의 일부와 유사하고, 다른 규칙의 결론부로 부터 값을 유도해야 될 경우
즉, 규칙들간의 다중 상호 의존관계를 갖는 경우 TAEMS 트리구조에서 생성된 연관 규칙들로
부터 값을 유도하는 과정



Input_Array[50]={Ca, Mg,Na,K,Cu,Zn,P,Fe,Mn,Cr,Se,B,Co,Ge,Mo,S,V,As,Be,Hg,Cd,Pb,Al,Sb, Zr}
Mid_Output_Array[20]={Ca/P,Na/K,Ca/K,Zn/Cu,Na/Mg,Ca/Mg,Fe/Cu,Ca/Pb,Fe/Pb,Fe/Hg, S/Pb}

그림 13 다중 퍼지 규칙베이스에서 규칙들 간의 상호의존관계

태와 FUZZY_INFER( )의 추론알고리즘(FUZZY_INFER_I, FUZZY_INFER_II)은 다음과 같다. 여기서 Dataset(Input Facts)는 규칙의 조건부에 입력되는 미네랄 또는 미네랄 성분비 값들을 말한다.

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 하

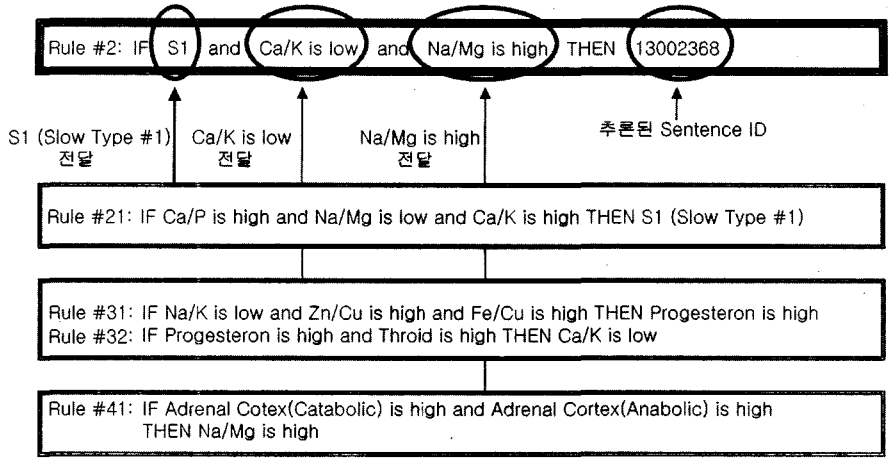


그림 14 다중 퍼지 규칙베이스에서 정보 추론 과정

는 이유는 TMA 미네랄 자료가 단순히 몇 개의 미네랄에 의해 분류되지 않고 매우 복잡한 연관관계를 통해 분석되고 측정되기 때문이다.

위 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 14와 같다. 만약 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 위의 그림에서와 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 스케쥴링 과정을 거쳐서 해당 ID 문장을 추론하게 된다. 각 규칙의 결론부가 상위레벨의 규칙의 각 조건부 문장들과 일치되는 것들을 찾아가면서 최종적으로 결론부의 문장 ID를 찾게 된다.

이로부터 추론된 해당 ID인 문장은 다음과 같다.

[ID: 13002368인 영어문장]

"Magnesium is essential for muscle relaxation, protein synthesis, nerve excitability and energy production on a cellular level.....(중략)... When magnesium levels are chronically low in the body, a tendency toward one or more of the following conditions will increase:

Noise Sensitivity, Tremors, Body Odor, Muscle Spasms, Irritability, Increased Blood Pressure....."

[ID: 13002368 번역된 한국어 문장]

"마그네슘은 세포레벨에서 근육이완, 단백질 합성, 신경흥분, 에너지 생산에 필수요소이다....(중략)... 만약 마그네슘 정도가 인체내에서 만성적으로 낮으면 다음과 같은 하나이상의 상태가 증가하게 된다.

소음 민감성, 흥분, 암내, 근육경련, 불안정, 혈압상승...."

따라서 환자의 다양한 미네랄 수치와 미네랄들간의 비에 따라 지능적으로 추론된 각 ID에 따른 영어문장들과 해당 문장들이 번역된 한국어 문장들이 데이타베이스로부터 추론되어 레포트형식을 갖추어서 결과지로써 환자와 의사에게 제공된다.

4. IMES 설계 및 구현

4.1 IMES 개요

본 논문에서 구현한 IMES는 미국에서 텍스트파일이 아니라 철저한 보안을 바탕으로 한 그래픽 파일로써 미네랄수치와 영문으로만 제공받던 기존의 국내 TMA 방법상의 문제점을 해결하기 위해, 그래픽 파일을 텍스트 파일화 하고 이들 미네랄로부터 유도되는 영어문장들과의 상관관계를 분석하여 데이타베이스화한 것이므로 TMA 분석된 미네랄 수치만 입력하면 국내에서도 미국에 의존하지 않고 자체적으로 TMA 분석 및 다양한 의학적 예측이 가능하도록 한 의학 전문가 시스템이다. IMES에는 미네랄로부터 분석된 다양한 형태의 한글화된 한국형 의학 정보 데이타베이스가 구축되어 있다. 따라서 IMES는 입력된 환자의 미네랄 수치에 의해 데이타베이스와 연동하여 각각의 조건을 검색하여 현재 환자의 상태 및 앞으로 발생할 상황들에 대한 예측을 다양하게 제공해준다.

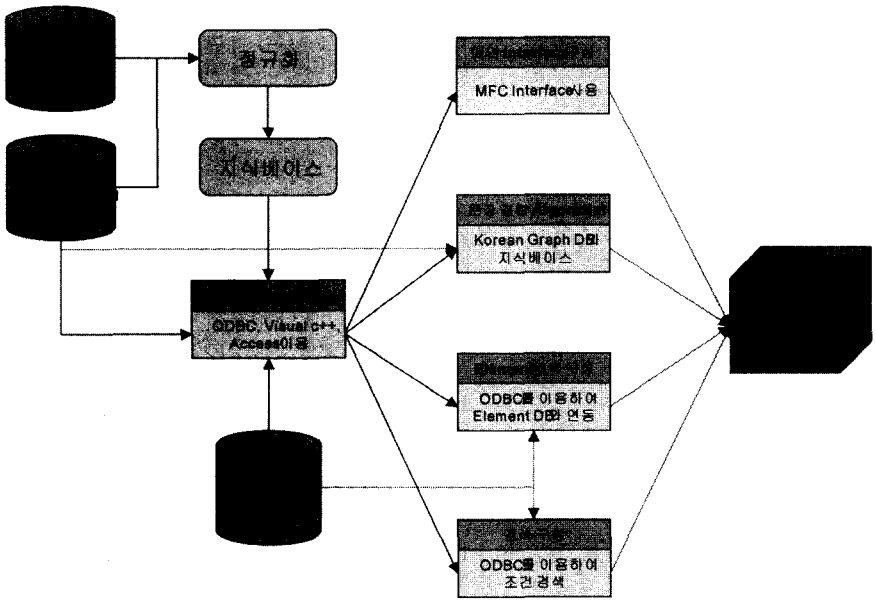


그림 15 IMES 시스템 구성도

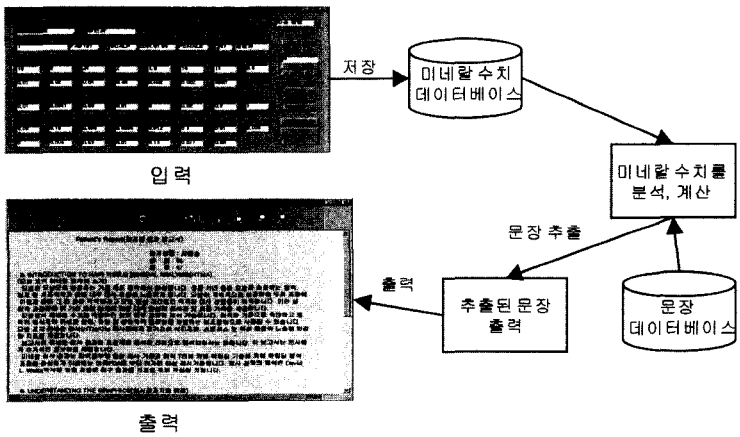


그림 16 IMES 시스템 I/O 및 구성요소간 연관성

본 논문에서 구현한 IMES의 구성도는 그림 15와 같다. 본 시스템은 MS Visual C++과 MS ACCESS를 이용하여 윈도우 환경에서 구현하였다. IMES에는 Analysis DB, Koream DB, Element DB가 있다.

IMES 시스템 I/O 및 구성 요소들간의 연관성은 그림 16과 같다.

IMES 시스템은 미네랄 수치입력 인터페이스와 포트를 출력해주는 부분으로 크게 나누어진다. 미네랄 수치를 입력하면 입력한 수치는 수치 데이타베이스에 저장되고 그 입력된 수치값으로 문장 데이타베이스에서 문장을 추출하게 된다.

데이타베이스는 환자에 대한 미네랄 수치와 그 환자의 인적사항과 기타 사항을 관리하는 TMA 테이블, 병원과 진료과, 의사 그리고 담당자를 관리하는 테이블 administrator 테이블, 환자의 재검사여부를 관리하는 ID_Management 테이블, 문장 수정과 탐색을 쉽게하기 위한 Sent_Name 테이블 등 4개로 구성하였다.

4.2 IMES 시스템 분석

4.2.1 분류방법에서의 오분류율 처리방법

분류방법의 적절성을 확인하는 절차 중의 하나는 오

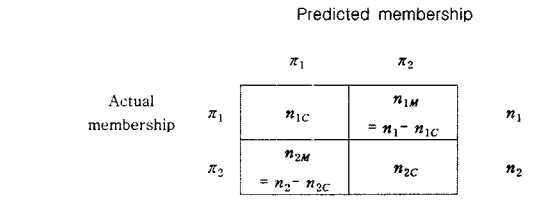
분류율(misclassification probability)을 계산하는 것이다. 본 연구에서 오분류율의 계산은 표본을 이용하여 다음과 같은 식 (1)에 의해 계산하였다.

$$ARE(actual\ error\ rate) = p_1 \int_{\hat{R}_2} f_1(x)dx + p_2 \int_{\hat{R}_1} f_{2(x)}dx \quad (1)$$

여기에서 $\hat{R}_1$과 $\hat{R}_2$는 표본의 크기에 의해 결정되는 분류지역(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 구성되어져 있으므로 표본정보만을 갖고 있는 본 연구에서는 모집단의 분포에 종속되지 않고 오분류율을 계산할 수 있어야 하기 때문에 APER(APparent Error Rate)을 이용하였다.

그림 17을 이용하여 APER = $\frac{n_{1M}+n_{2M}}{n_1+n_2}$로 계산되며 이는 오분류된 비율을 나타낸다. 본 논문에서 오분류율 측정 결과 8가지 대사타입(빠른 대사, 느린 대사)에 대해서는 오분류율이 거의 0에 가까웠으나 표본의 크기가 작은 미네랄간의 관계 등에서는 그림 18 분포율과 같은 오분류가 발생하였는데, 표본크기가 200 이상일 경우 오분류율은 거의 0에 가까운 값을 나타내었다.



여기에서 n_{1C} = 실제 π_1에 속하면서 π_1에 분류된 수
n_{1M} = 실제 π_1에 속하면서 π_2로 잘못 분류된 수
n_{2C} = 실제 π_2에 속하면서 π_2에 분류된 수
n_{2M} = 실제 π_2에 속하면서 π_1으로 잘못 분류된 수

그림 17 오분류율 측정방법

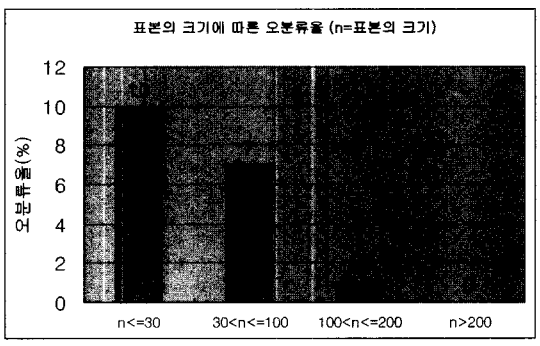


그림 18 오분류율

4.2.2 시스템의 효율성

표 4에 나타난 바와 같이 다중 퍼지 규칙베이스를 구축하였을 경우 단일 퍼지 규칙베이스를 구축하는 것보다 시스템의 효율성이 향상되었다. 이것은 표본크기가 2000이고 오분류율이 0이었다는 가정하에 측정된 결과이다.

표 4 IMES 시스템의 효율성

구분 / 비교항목	다중 퍼지규칙 베이스 (제안한 방법)	단일 퍼지규칙 베이스 (기존의 방법)
평균 처리속도(sec)	0.1sec	0.5sec
결과의 정확성(%)	92%	68%

평균 처리속도는 입력창에서 미네랄 수치를 입력하여 영어 및 번역된 문장을 찾아내는 속도를 측정한 것으로, 기존의 방법은 규칙의 입력부분의 비교횟수가 많아 처리속도가 느린 반면 제안한 방법은 규칙의 조건부 비교횟수를 줄여주어 결론부를 재입력함으로써 평균 처리속도가 약 5배 정도 향상되었다.

결과의 정확성은 2001.7-2002.2까지 7개월간 국내에 보관된 환자의 임상자료 약 2000건에 대해 측정한 것으로, 시스템에서 전체적으로 한국어로 번역되어 출력되는 최종 결과지가 미국에서 보내오는 영문으로 된 결과지와 일치하는 정도를 측정한 것이다. 환자 한건당 결과지에 포함해서 출력해야 할 총 50개의 문장들 중 단일 퍼지규칙 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 IMES 시스템의 데이타베이스에 저장된 총 영문 종류는 각 대사 및 미네랄 타입별로 약 700개 정도이다.

4.2.3 시스템의 활용성

다음 그림 19와 같이 IMES 시스템을 사용한 경우와 사용안한 경우로 나누어서 시스템의 활용성에 관한 실험을 하였다.

분석을 위해 이 프로그램을 사용할 (주)한국 티이아이를 대상으로 2002.3월부터 2002.6까지 임상실험 결과를 중심으로 실험을 하였다. 이 회사에서는 환자로부터 모발검사를 실시한 후 미국 TEI에 의뢰하여 결과를 받은 후 한국에서 프린트하여 환자와 담당 의사에게 돌려주게 되어 있다. 환자 1건당 이러한 전 과정을 거치는데 2002.3월 이전에는 보통 50분 정도의 시간을 소요하였는데, 실험결과 IMES를 사용할 경우 약 7분 정도가 소요되어 IMES를 사용하지 않은 경우보다 86% 일의능률을 가져올 수 있었다.

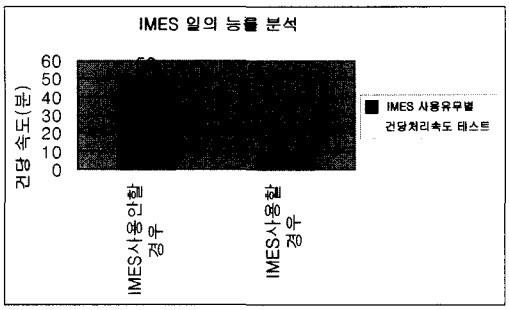


그림 19 IMES 시스템의 일의능률 분석

이상으로부터 이 회사의 IMES 시스템에 대한 활용도 대비 일의 능률에 대한 만족도는 다음 그림 20과 같다.

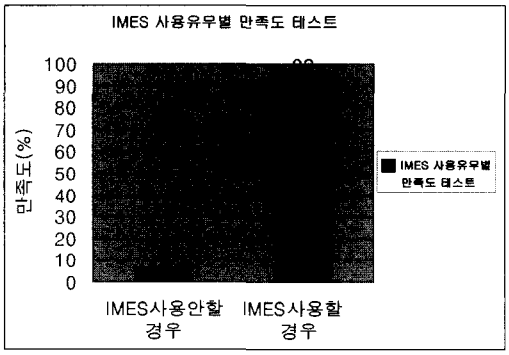


그림 20 IMES 시스템의 만족도 분석

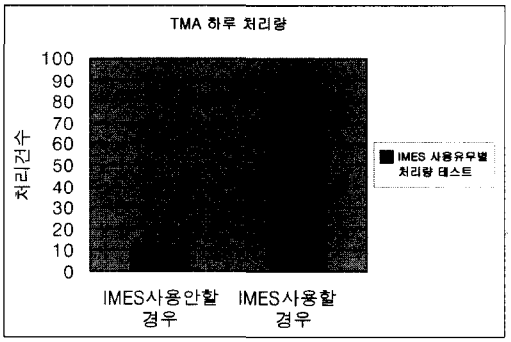


그림 21 IMES 시스템의 사용여부에 따른 TMA 처리량 분석

이 그림에서 만족도란 활용도를 100으로 했을 때의 일의 능률 값의 비를 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 IMES가 없는 상태에서의 처리속도이며 3은 최소 기본 처리속도를 말하고, 3은 '만족도 100'을, 50은 '만족도 0'을 의미한다.

IMES 만족도 = 100 * (Max-소요시간) / (Max-Min) (2)

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

IMES 사용여부에 따른 하루 TMA 처리량(10시간 기준)은 그림 21과 같다.

4.3.4 시스템 개발에 대한 기대효과

IMES 시스템 및 의학 정보 데이타베이스는 다음과 같은 기대효과를 갖는다.

- 한국과 유사한 동남아 지역 의학 관련 기관들에게 유용(수출가능)
- 자신의 영양불균형을 예방하여 건강을 유지하려는 사람들에게 정보제공
- 한국형 국민 건강지료로부터 추가 의료 및 바이오 연구분야 발달 가능성
- 현대의학으로 치료가 불가능한 분야의 치료개발 기술 제공 가능성
- 의료계에 질높은 의료서비스 제공

5. 결 론

본 논문은 TMA(모발분석)를 위한 한국적 데이타베이스 구축 및 지능적 의학 전문가 시스템 개발에 관한 것으로 기존의 국내 TMA 방법에서의 다음과 같은 문제점을 해결하기 위해 연구하였다.

첫째, TMA 분석기기는 있으나 분석결과를 해석할 수 있는 의학 전문가 프로그램이 개발되어 있지 않다. 둘째, 한국에 미국과 같은 의학정보 데이타베이스가 없으므로 분석방법 개발이 필요하다. 셋째, 미국에서 분석되어 오는 TMA 결과가 영문이며 보안위주 파일이므로 가독성 및 재활용성이 떨어진다. 넷째, TMA 관련 데이타베이스 구축 수준이 매우 낮으므로 TMA 분석 및 의료서비스를 위해 매번 미국에 의뢰해야 하므로 심각한 외화낭비를 초래한다. 다섯째, TMA 결과가 서구식 생활패턴에서 비롯된 데이타베이스로부터 구축된 것이므로 보다 분석적인 한국형 데이타베이스구축 및 TMA 결과분석이 필요하다.

본 논문에서는 이러한 문제점을 개선하기 위해 다음과 같은 방법을 제안하였다. 첫째, 통계적 TMA 자료분석을 위해 다단계 분류방법을 활용하여 다중으로 영향을 주는 요인들에 대하여 분류하였다. 둘째, 다중 퍼지 규칙베이스를 구축하여 지능적인 퍼지추론 방법에 의해 미네랄들을 찾아내고 관련성을 연구하였다. 셋째, 이러

한 방법을 이용하여 지능적 의학 전문가 시스템(IMES: Intelligent Medical Expert System)을 구축하였다. 넷째, 한국형 TMA 데이타베이스를 구축하였고 이를 위해 한글화 작업을 병행하였다.

현재 개발된 IMES 시스템은 (주)한국 티이아이에서 사용 중에 있으며 국내 임상실험한 자료가 미국 TEI에 비해 작고(약 2000 여건 정도) 개발기간도 7개월 정도였지만 업무능률이 IMES 시스템을 사용하지 않았을 경우보다 86%이상 향상되었고, 사용자 만족도가 92%정도 되었고, 시간당 처리하는 건수도 기존에 비해 약 7.5배 이상 증가하게 되었으며 결과의 정확도도 약 92%정도로 측정되었고, 시스템의 평균 처리속도도 약 5배이상 증가하였다.

IMES 시스템과 의학 정보 데이타베이스는 지금의 추세로 보아 향후 2년내 국내시장 성장률 660%, 세계시장 성장률 300%를 가져올 것이며, 향후 3년내 200억원의 수입대체효과와 200억원의 수출효과를 가져올 것으로 기대한다.

앞으로 연구해야할 과제는 좀더 많은 임상결과로부터 시스템을 개선하는 일이고 지능적 추론 방법과 다단계 통계적 분류방법과의 연관성에 관해 좀더 연구하여 시스템의 오류율을 최소화해야 하는 일이며 웹상에서 서비스할 수 있도록 시스템을 확장하는 일이다. 또한 미국 TEI로부터 얻은 식단은 한국의 현실과 동떨어진 경우가 많으므로 실제적으로 도움이 되는 한국형 식단개발도 계속 연구해야 할 과제이며, 앞으로 신약연구개발이나 국민영양지표 표준화 등의 사업에 계속적으로 응용해야 할 것이다.

참 고 문 헌

[1] Watts, D,.L., Trace Elements and Other Essential Nutrients-Clinical Application of Tissue Mineral Analysis, 2nd Writer's BLOCK edition, 1997.

[2] 의계신문(The medical World News), 1247호, 2001. 1.22일자.

[3] http://www.traceelements.com.

[4] http://www.teikorea.com.

[5] http://www.medinex.com.

[6] Strain, W. H. and Pories, W. J. and Flynn, A. and Hill, O. A., Trace Element Nutriture and Metabolism Through Head Hair Analysis. Trace Substances in Environmental Health. Hemphill, D., Ed. Univ. Mo. Press, Columbia. 1972.

[7] Baumgartner, W. A. and Hill, V.A. and Blahd, W. H., "Hair Analysis for Drugs of Abuse," Journal of Forensic Science Vol 34. No 6. 1989.

[8] Vir etal, "Serum and hair concentrations of Copper During Pregnancy," American Journal of Clinical Nutrition Vol. 34, 1981.

[9] Laker, M. "On Determining Trace Element Levels," Man The Uses of Blood and Hair, Lancet 11, 1982.

[10] 장수익, 김경곤, 이복기, 김형준, 유수헌, 강희철, 윤방부, "당뇨환자에서의 모발내 미네랄의 함량연구", 최신의학 45권 5호 별책. 2002.

[11] 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.

[12] 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.

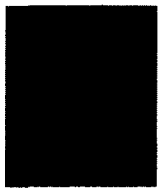
[13] Johnson, R. A. and Wichern, D. W, Applied Multivariate Statistical Analysis 3rd edition, Prentice Hall, 1992.

[14] 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.

[15] 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.

[16] Cho. Y.I, "An Effective Hybrid Methodology of Neuro-Fuzzy Systems," *International ICSC Congress Computational Intelligent Methods and Applications*, USA, NY, pp.69~74, June 22-25, 1999.

[17] Graham, J. R. and Decker, K.S., "Towards a Distributed Environment-Centered Agent Framework," Umass Technical Report, 1998.



조 영 임

1987년 고려대학교 생물학과 졸업. 1988년 고려대학교 컴퓨터학과 졸업. 1990년 동 대학교 대학원 컴퓨터학과 석사. 1994년 동 대학교 대학원 컴퓨터학과 박사 1996년 평택대학교 컴퓨터과학과 조교수 1999년~2000년 Univ. of Massachusetts at Amherst, post-doc. 2001년~현재 한국공학교육학회 편집위원. 2003년~현재 한국퍼지 및 지능시스템학회 홍보이사 겸 편집위원. 관심분야는 인공지능, 에이전트, 뉴로-퍼지 시스템, 가상도서관

한 근 식

1984년 고려대학교 통계학과 졸업. 1990년 Iowa State University 통계학 석사 1993년 Oklahoma State University 통계학 박사. 1994년 한신대학교 정보시스템 공학과 교수. 관심분야는 표본이론, 정보시스템