

전문가 시스템 구성을 위한 범용 추론 시스템 - CIS

(CIS - A General Purpose Inference System for Building Expert System)

朴 鍾 勛*, 崔 宗 秀*

(Jong Hoon Park and Jong Soo Choi)

要 約

본 논문은 전문가 시스템 구성을 위한 새로운 범용 추론 알고리즘과 그에 의해 구성된 CIS (Condition Interpret System)에 대해 논한다.

본 논문의 목적은 과거의 전문가 시스템들에 나타났던 지식 표현의 단편성과 난해한 추론 기법 등을 해결하기 위해 좀 더 자연스럽게 일관성 있는 지식표현 방법과 그에 따르는 유연한 추론 기법을 제안하는 데 있다.

CIS는 사용자의 몇몇 상황 설명을 토대로 대화를 통해 현재 일어나고 있는 모든 상황들을 추론해낼 수 있는 시스템으로서, 고혈압과 그 합병증에 관한 매우 간단한 의료 지식 데이터에 의해 시험 구동되었다. 그 결과, 본 논문의 지식 표현과 추론 기법에 의해 훌륭한 전문가 시스템 구성이 가능할 것으로 보이므로 이에 보고한다.

Abstract

In this paper, we study a (condition interpret system CCIS) which is a general purpose inference system for building an expert system. Previously, most expert systems used fragmentary knowledge-representation and inexact inferencing methods. To solve the problems of those systems we suggest an algorithm for more flexible deduction and a method of representing knowledge more naturally and consistently compared to those of other expert systems. The CIS is a system which can infer all the current situations from communication with the user based on user discription of some situation. We used some medical data on hypertension and its complication to simulate the system and to prove the effectiveness of the system. The results show that with the proposed method one can realize an attractive expert system.

I. 序 論

최근 인공지능이라는 학문의 힘을 빌어 종전의 컴퓨터 시스템들과는 비교도 될 수 없는 강력한 능력을 가진 전문가 시스템 (expert system)들에 관한 연구

가 급속화되고 있다. 이는 매우 빠른 속도로 발전해 가고 있는 각 분야의 전문 지식들을 그 분야의 전문가들이 모두 수용하기 힘들 뿐 아니라 문헌에 의해 그를 충족시키도 한계에 도달해 가고 있기 때문에, 그를 컴퓨터의 도움에 의해 타개해 보고자 하는 것이다. 컴퓨터는 “한번 입력된 것은 절대 잊지 않는다” 그리고 “절대 선입관이 없으며 추론에 실수가 없다”고 하는 커다란 장점이 있어, 이러한 시스템들이 실용화되었을 때

*正會員, 中央大學校 電子工學科
(Dept. of Elec. Eng., Chung-Ang Univ.)
接受日字: 1985年 10月 23日

는 인간에게 크나큰 도움을 줄 수 있을 것이다.

전문가 시스템의 연구는 1960년대 중반에서 시작하여 DENDRAL과 MACSYMA가 처음 만들어졌고, 그 후 의료 자문 시스템들, 즉 MYCIN, CASNET, PIP, INTERNIST 등이 지식 표현과 추론 과정에 대한 체계를 확립하면서 발전되었다.^{1,2,3,4,7,8)} 그리고 최근에는 범용성을 가진 전문가 시스템개발 tool, 예를 들면 EXPERT, EMYCIN, MECS-AI, UNIT 등에 관심이 집중되고 있으며 그에 의해 만들어져 여러 분야에서 실용화 단계에까지 이른 시스템들도 있다.^{1,5,6)}

이렇게 발전되어 온 전문가 시스템들에 있어 가장 중요한 생점은 지식 표현과 추론 기법이라 할 수 있겠다.^{1,2,5)} 그 지식 표현법으로 production rule, semantic net, frame 등이 있으며 추론 기법으로는 forward reasoning, backward reasoning과 그 융합형 등이 쓰이고 있다.^{1,2,7,8)}

1. 인간의 사고과정과 CIS의 기본개념

일반적으로 인공지능 분야의 많은 연구자들은 지식 표현법이나 추론 기법을 창출해내기 위해 인간의 사고과정에서 힌트를 얻어내고 있다. 그 이유는 인간의 사고과정이 가장 이상적 추론기능을 수행하고 있으며 그 시스템들을 사용할 주인도 역시 인간이기 때문이다. 그러므로 전문가 시스템들도 인간의 사고과정과 비슷한 추론을 행함으로써 인간에게 친근한 시스템으로 만들어지는 것이 바람직하다.

그러면 과연 우리 인간들은 어떤 형태로써 지식을 구축하며 그를 어떻게 추론해내는 것일까?

우선 우리가 가지고 있는 지식이라고 하면 경험적으로 얻어진 단편적 지식과 학습에 의한 구조적 지식으로 분류할 수 있겠다.^{1,5,6)} 또한 이러한 지식들은 인간의 사고방식에 의해 몇몇 연관관계를 유지하며 복잡하게 얽혀 있다고 생각된다.

그 첫번째 연관관계의 예가 종속(hierarchy)에 관한 개념이다. 인간은 모든 개체 혹은 상황(situation)을 좀 더 넓은 그리고 좀 더 일반적인 곳에 종속시키려는 경향이 있다. 이러한 관념은 온갖 동물들을 복잡하게 나누어서 유목적(類目的)으로 분류해 놓은 것을 보더라도 단적으로 나타난다.

둘째로 인간은 많은 상황(situation)들을 인과관계에 의해 묶어 놓고 있음도 볼 수 있다. 우리는 어떤 상황이 일어나면 반드시 그 원인을 규명하려는 경향이 있다. 또한 이전의 경험에 의해 결성되어 있었던 인과관계에 의해 앞으로의 일들을 추측해내기도 한다.

마지막으로 특성과 연상에 관한 관념을 인간은 가지고 있다. 인간은 개체나 상황들의 특성들을 잘 생

각해 두었다가는 다음에 그를 다시 확인할 때 그 특성들을 이용한다. 그리고 그 특성들중 하나를 우연히 보게되면 즉시 그에 해당하는 개체나 상황을 자연스럽게 연상해내는 것이다.

위에서 설명한 관계 외에도 몇몇 관계들이 있을 수 있겠다.

이러한 사고방식을 볼 때 인간의 지식은 개체나 상황(situation)들이 복잡하게 연관지어져 얽혀 있을 것이다. 이렇듯 복잡하게 연결된 지식들을 자연스럽게 컴퓨터에 이식시키려면 매우 강력한 체계를 갖춘 연관관계 표현법이 필요하다.

인간의 지식을 IF~THEN의 단편적인 형식에 의해 표현하고 있는 rule-based system들은 이러한 점에서 취약점을 안고 있을 수 밖에 없는 바 그 system들이 간결하고 편리한 추론 메카니즘을 가지고 있기는 하나 인간의 복잡한 지식체계를 표현하는 데는 한계가 있다. 한편 일련의 frame-based system들에서 인간 지식의 연관관계를 표현하고 추론에 이용코자 연구되어 오고 있으나 이는 난해하고 복잡한 지식표현과 추론기법 때문에 극히 제한된 영역에서 밖에는 환영을 받고 있지 못하다.^{1,2,5,6)}

2. CIS의 기본 짜임새

이러한 문제점 해결을 위해서 본 논문에서는 일상 생활에서 일어나는 상황들을 기본 단위(node)로 하여 위에서 설명했던 가장 중요한 세가지의 연관관계를 자연스럽게 표현할 서술적 지식표현법과 유연한 추론 알고리즘을 소개하고 있다.

본 시스템의 지식표현은 기본 단위인 상황(situation)들에 대해 각각 7개의 일반상황지식(연관관계기술), 즉 소속과 세분, 원인과 결과, 특성과 연상 그리고 대응책을 서술적으로 기술해준다. 이렇게 함에 의해 최초의 원인이 될 수 있는 상황부터 최후의 결과가 될 수 있는 상황까지가 인과관계의 계통을 이루면서 이어지며 최고의 포괄적 상황부터 최하의 가장 세분된 상황까지 종속의 계통이 이루어진다.^{1,10)} 또한 각 node가 될 상황들에 대해 각각 그 특성과, 연상되는 상황들이 기술된다. 이러한 관계들은 추론과정에서 각기 나름대로의 의미를 가지며 이용된다.

본 시스템의 추론과정은 가설의 설정, 선택, 검증을 번갈아 행함에 의해 이루어지며 검증과정에는 가설의 소속추적, 특성평가, 원인규명 과정이 포함된다. 이러한 과정은 현재 일어나고 있는 모든 상황들을 찾아낼 때까지 계속된다.

그런데 본 시스템의 추론과정은 일반적 시스템에서의 특성평가에 의한 가설 검증법에 소속추적과 원인규

명 과정을 포함시킴으로써 종전의 시스템들에 비해 좀 더 완벽한 추론을 해낼 수 있게 된다. 또한 확실한 상황에 대한 연관관계들에 의해 직접 가설을 설정하고 그에 각각 특정한 값(weight)을 주고 그를 가설 검증 때마다 평가 수정하여 그 값에 의해 가설을 선택하기 때문에 종전의 다른 시스템들에 비해 가설의 설정과 선택방법이 훨씬 용이하다는 장점을 가진다. 가설의 설정과 선택방법이 간결해야만 앞으로 방대한 분량의 전문 지식을 운용해야 할 경우 추론엔진이 위력을 발휘할 수 있을 것이다.

이러한 몇몇 장점들은 각 상황들에 대한 강력한 연관관계 기술에 기인하는 것이며 본 시스템에서의 지식표현의 유용성을 입증할 수 있는 증거이다.

물론 본 시스템의 단점들도 몇가지 나타나고 있는데 이는 결론부에서 언급하기로 하고 다음의 본문에서는 본 시스템의 지식표현법과 추론기법을 상세히 논하고자 한다.

II. 지식 표현(Knowledge-Representation)

본 시스템의 모든 지식은 상황(situation)을 단위로 LISP의 List구조로써 표현되고 운용된다.¹¹⁾ 여기서 상황이란 단일의 사건 또는 현상을 말하며, 예를 들면 “그녀는 아름답다”라든지, “나는 집에 간다” 또는 “그는 딸에게 인형을 주었다” 등이다. 이때 “그는 현기증을 느끼며 두통이 있다”와 같이 여러가지 상황이 접속되어 한개의 상황으로 이루어진 것은 몇개의 단일 상황으로 나누어 표현한다.

1. 상황 표현

먼저 한개의 상황이 존재할 때 그 상황 자체를 표현할 적당한 방법이 있어야 하므로 다음과 같이 상황을 표현한다. 모든 상황은 그 상황의 중심이 되는 술어(predicate)와 그 술어를 한정해 주는 한정어(specifier)로 구성되는 상황 pairlist로써 표현된다.

(술어 · 한정어) : 상황 pairlist

이 때 이들 술어와 한정어는 각각 다시 상황 pairlist로 구성될 수 있다.

“그녀는 아름답다” → (아름답다, 그녀)

“환자의 질병은 고혈압이다” → ((이다, 고혈압), (질병, 환자))

“나는 집에 간다” → ((간다, 집), 나)

“그는 딸에게 인형을 주었다” → (((주었다, 인형), 딸), 그)

여기서 각 단어들 간에 혹은 list간에 한정 관계가 이루어지는데 두개 이상의 단어들 간에 한정 관계가 아닌 동등 관계가 주어지는 경우에는 술어로서 특수

문자인 “*”을 쓰고 그 한정어는 동등관계를 이루는 단어들로 이루어진 list로 한다.

“버스와 택시가 충돌했다” → (충돌했다, (*, (버스 택시)))

“2와 4와 9를 더한다” → (더한다, (*, (2 4 9)))

우리가 취급하고자 하는 상황은 비교적 간단한 구조를 가지고 있으며 복잡한 수식어 등의 사용은 금지하고 있기 때문에 이러한 방법에 의해서 모든 상황이 표현 가능하게 되고 상황 pairlist의 한글과의 interface 구성도 매우 용이하게 이루어질 수 있다.¹¹⁾

이렇게 표현된 각 상황들은 지식의 기본단위이고 추론과정에는 전혀 영향을 주지 않는다.

2. 상황지식 표현

위에서 각 상황들 자체를 기술하는 방법을 설명했는데 본절에서는 그 각 상황에 대한 지식을 표현하는 방법에 대해 논한다. 본 시스템에서는 다음과 같이 각 상황에 대한 지식을 표현한다.

가) 일반상황지식 표현 (long-term memory)

일반상황지식이란 전문가들이 가지고 있는 일반적인 지식을 의미한다. 이는 시스템 구동시 변경되지 않는 고정된 지식이다. 본 시스템에서는 일반상황지식의 표현에 있어서 각 상황에 대해 다음과 같은 7가지의 연관관계를 기술해 준다.

[원인-결과 관계]

① 원인관계: 그 상황에 대한 원인이 되는 상황들이 이에 해당한다. 예를 들어 “환자의 질병이 고혈압이다”라는 상황에 대해 그 원인이되는 “비만하다” 이라든지 “동맥경화이다” 등이 이 부분에 기술된다.

② 결과관계: 그 상황에 의해 결과로서 나타날 수 있는 상황들을 말하며 예를 들면 “고혈압이다”라는 상황에 대해 그 결과로써 “뇌질환”, “심장질환”, “신장질환” 등을 포함하며 이들 각각은 Dependency가 없다고 가정하였다.

[소속-세분 관계]

③ 소속관계: 그 상황이 일어나고 있는 배경이 되며 좀 더 일반적인 의미를 가지는 상황들이다. 이 경우는 “뇌출혈이다”라는 상황이라 하면 그의 상위개념인 “뇌질환이다”등을 말한다.

④ 세분관계: 그 상황에서 좀 더 세분되어진 상황들이다. 즉 고혈압은 신장성, 본태성, 그리고 내분비성으로 나누어지는데 바로 “신장성고혈압이다” 등등의 상황들을 여기에 포함시킨다.

[특성-연상 관계]

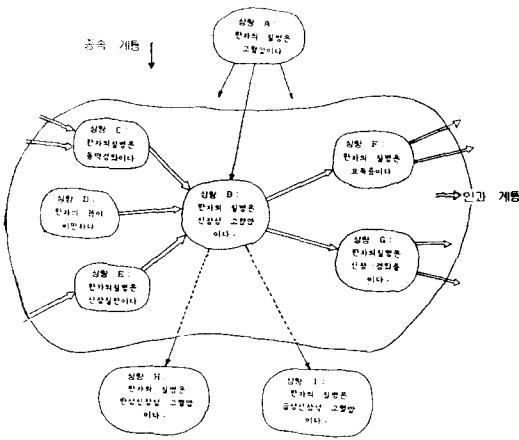
⑤ 특성관계: 그 상황의 특성으로 나타나는 상황들이다.

씨 시스템 구동 도중에 발체한 것이다.

다) 인과계통과 종속계통

CIS에서 일반 상황 지식 표현에 의해 인과 계통과 종속 계통이 자연스럽게 이루어지게 된다.¹¹⁾

인과 계통은 인과 관계를 가지고 상황들 간에 횡적으로 이루어지는 계통이다. 한편 종속 계통은 소속과 세분 관계에 의해 종속의 관계를 가진 상황들 간을 종적으로 이어주는 계통이다. 그러므로 인과 계통과 종속 계통에 의해 거의 모든 상황들은 횡과 종의 삼차원적 구조로 얽히어 있게 된다(그림3 참조).



- * 상황 C, D, E, B, F, G가 인과 계통을 이룬다.
- * 상황 A, B, H, I가 종속 계통을 이룬다.

그림 3. 인과계통과 종속계통
Fig. 3. Causal and hierarchical genealogy.

이렇게 상황들 간에 이루어져 있는 구조는 추론과정에서 그 계통들을 확인함으로써 가설검증의 효율성과 정당성을 가지도록 해주는 역할을 한다.

Ⅲ. 추 론

1. 추론 알고리즘

본 시스템의 추론은 다음과 같이 이루어진다.

다음의 그림 4에서 본 시스템에서의 기본 골격이 되는 추론 엔진은 가설설정, 가설선택, 가설검증의 세과정이며 나머지의 부수적 과정은 다른 분야에서 사용된다면 여러가지 형태로 바뀌어 운용될 수 있을 것이다. 여기서는 편이상 고혈압에 대해 시험구동에 이용되었던 시스템에 준해서 설명한다.

가) 시작과 초기 데이터 입력

시스템이 구동되면 기본적으로 필요한 데이터를 시

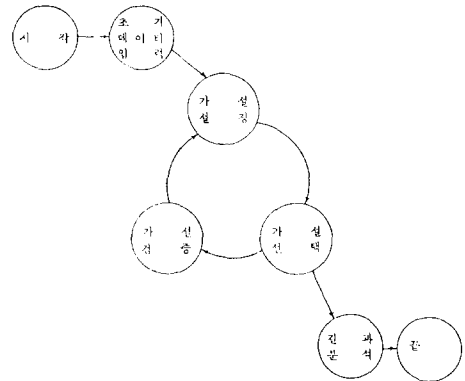


그림 4. 추론과정 개략도
Fig. 4. Brief diagram of inferencing procedure.

스템이 사용자에게 요구한다. 환자의 이름, 나이, 성별, 과거 병력, 현재 증상들을 묻고 그를 토대로 가설을 세운 뒤 추론을 시작한다.

나) 가설 선택

가설에 포함된 상황들 중에서 SPF(특수 가능성 지수)가 가장 큰 것을 선택하여 검증할 가설로 한다. 이는 지금 있을 수 있다고 추측되는 상황들 중 그 가능성이 가장 큰 것을 먼저 검증하기 위해 선택한다는 의미이며 그는 가설들 각각이 가지는 SPF값에 의해 간단히 선택되어질 수 있다.

다) 가설 검증

가설의 검증 과정은 다음과 같이 이루어진다.

[과정1]: 가설의 종속 계통을 추적하여 확인한다.

가설의 소속 관계 상황들을 추적하여 모은 후 이를 위로부터 하나씩 차례로 검증한다. 만일 그 도중에 한 상황이 부정 또는 미정인 경우에는 그 밑의 상황들은 검증할 가치를 잃게 된다. 소속 관계 상황이란 그 상황이 있기 위한 배경, 즉 필수 조건이기 때문에 그것이 확정되지 않으면 지금 검증하고자 하는 상황은 평가해 볼 필요도 없게 되는 것이다.

[과정2]: 가설의 특성들을 검증한다.

특성관계 상황이란 가설에 대한 특성들이기 때문에 이를 검증해 보아야 실제로 그것이 사실인가를 알게 되는 것이다. 그 방법은 가설의 특성관계에 속한 상황들을 모두 평가해보아 CF(확실성지수)가 임계값인 70을 넘으면 그 상황을 사실이라 인정하여 확정집합에 두며 그렇지 않으면 부정 혹은 미정집합에 둔다. 이때 시스템의 힘으로 검증할 수 없는 상황, 즉 특성관계가 주어지지 않은 경우(주로 시스템에서는 환자의 증상들이 이에 해당)에는 사용자에게 질문을 하여

알아낸다. 질문에의 대답은 -100부터 100까지의 숫자로 하며 -100~-70은 “그렇지 않음”, -69~70은 “잘 모르겠음”, 71~100은 “사실임”을 나타낸다.

한편, 위에서 언급한 임계값 70은 시험 구동해본 결과 가장 알맞다고 생각된 값으로 채택된 것이며, 이는 시스템이 실제 구현될 때 다시 정해질 수 있다.

[과정3]: 가설의 인과계통을 확인한다. 가설의 원인 관계에 속한 상황들을 확인해 보아 모두가 부정 혹은 미정일 때에는 그 가설이 특성 검증에 의해 확정으로 판명되었더라도 확정을 보류하게 되며 그의 원인들 중 아직 검증되지 않은 상황들을 가설집합에 편입시키고 그 원인이 검증후 확정이 되면 다시 검증하여 확정시킨다. 이때 원인관계 상황이 없는 가설은 예외로 취급되어 원인 불명의 사실일 수도 있다는 가정 하에 무조건 확정시킨다. 여기서 원인을 확인하는 이유는 원인이 확실치 않은 사실은 믿을 수가 없다는 생각 때문이다.

라) 가설설정

가설이 검증후 확정이 되면 그 가설과 연관관계를 가진 상황들에 대해 SPF(특수가능성지수)를 계산한 후 그 값을 가지고 가설집합에 포함시킨다. 이때 이전에 있던 각 가설들에 대해서도 SPF(특수가능성지수)를 검사하고 수정하며 그 값이 음이 된 것은 가설로부터 제외시킨다. 이렇듯 검증된 상황과 연관관계를 가진 상황들을 그대로 가설로 취하면 되므로 다른 시스템에서 어려운 문제가 되는 가설설정이 쉽게 이루어지며 한개의 가설이 검증될 때마다 남은 가설들의 가능성이 변화되기 때문에 좀 더 효율적인 추론이 가능하다.

마) 결과분석 및 끝

해석할 가설이 남아 있지 않으면 시스템은 추론을 끝내고 결과분석에 들어간다. 먼저 추론과정에서 확정되었던 상황들을 질병과 증상으로 구분하고 증상들을 출력해 준다음, 질병의 원인과 소속 등 연관관계가 있는 상황들을 알려주게 된다. 또한 확정된 상황들의 아직 나타나지 않은 결과관계 상황들에 의해 앞으로 일어날 수 있는 상황들을 추측해낼 수도 있다. 즉 한가지 상황이 일어났을 때 그 결과로서 나타나곤 하던 상황들이 아직 나타나지 않았다면 이는 앞으로 일어날 가능성이 상당히 많다는 뜻이 되기 때문이다.

이렇게 현재 일어나고 있는 상황들의 설명과 앞으로 나타날 수 있는 상황들에 대한 자문을 해주고나서 시스템은 모든 동작을 끝내게 된다.

2. 지식의 불확실성 처리

전문가 시스템에서의 불확실성 처리는 매우 중요한

요소로서 지식표현과 추론 과정의 특성에 따라 알맞는 방법으로 수행되어야 한다. CIS에서는 추론과정에서 지식의 불확실성을 처리하기 위해서 검증 전의 상황에 대한 가능성지수와 검증 후의 상황에 부여되는 확실성지수로 나누어 운용한다.

가) 가능성지수(PF:Possibility Factor)

가능성지수는 추론과정에서 아직 검증되지 않은 상황들의 예상되는 가능성을 나타내주는 지수이며 둘로 나누어진다.

○일반가능성지수(General Possibility Factor)

일반가능성지수는 일반상황지식에서 연관관계의 내용이 되는 상황들이 가진 값으로서 일반지식을 입력할 때 전문가가 지정하며 일반상황지식(long-term memory)의 일부분이므로 추론시 값이 변하지 않는다. 이는 가설 검증시 특성들의 중요도를 나타내주거나 어떤 상황이 가설로 설정될 때 그 가능성을 표시해준다. 이 값은 -100부터 100까지의 숫자로 표시되며 70~100은 그 상황의 가능성이 거의 확실함”을 나타내고 -100~-70은 가능성이 “거의 희박함”을, 그 중간값을 가지면 가능성은 있으나 확실한 것은 “해석 후에야 알 수 있음”을 의미한다.

○특수가능성지수(Special Possibility Factor)

특수가능성지수는 특수상황지식(short-term memory) 중 가설들이 가지고 있는 예상되는 가능성을 나타내주는 지수이다. 이 지수는 검증 전의 가설들에게만 주어지고 검증할 가설을 선택하는 데 사용되며 검증이 끝나면 즉시 CF(확실성지수)로 교체된다. 또한 선택된 가설의 검증이 끝날 때마다 가설집합에 속한 상황, 즉 가설들 모두는 각각 그 SPF값이 갱신되는데 그 과정은 다음과 같다.

어떤 한 시점에서의 가설집합을 H_{n-1} 이라 하면

$$H_{n-1} = \{h_1(SPF_{h_1, n-1}), h_2(SPF_{h_2, n-1}), \dots, h_m(SPF_{h_m, n-1})\}$$

여기서 $h_1 \sim h_m$ 까지의 가설중 SPF가 가장 큰 것이 검증될 가설로 선택된다. 이를 h_{MAX} 라 하자. h_{MAX} 를 검증하여 그 CF(확실성지수)가 $CF_{h_{MAX}, n}$ 으로 나타났다면 나머지 가설들 각각의 SPF값은 다음과 같이 수정된다.

(i) 가설 h_m 이 h_{MAX} 와 연관관계가 있을 때

$$SPF_{h_m, n} = \frac{SPF_{h_m, n-1} + (CF_{h_{MAX}, n} * GPF[h_m/h_{MAX}]) / 100}{2}$$

여기서 $GPF[h_m/h_{MAX}]$ 는 h_{MAX} 의 일반상황 지식중 h_m 의 가능성지수, 즉 h_{MAX} 에 대한 h_m 의 연관관계의 중요도이다. 또한 $SPF_{h_m, n-1}$ 은 가설 h_m 이 애초에 가지고 있던 SPF값이며 $SPF_{h_m, n}$ 은 h_m 의 SPF 갱신작

업이 완료된 후의 값이다. $SPF_{h_m \cdot n}$ 은 h_{MAX} 의 CF값에 의해 증가 혹은 감소한다.

(ii) 가설 h_m 이 h_{MAX} 와 연관관계가 없을 때
 윗식에서 $GPF[h_m/h_{MAX}]$ 가 0 (즉, 연관관계가 없음)이므로 윗식은 다음과 같이 변형된다.

$$SPF_{h_m \cdot n} = SPF_{h_m \cdot n} / 2$$

즉 가설 h_m 의 SPF값은 반감된다.

이렇게 해서 h_{MAX} 의 검증후 가설집합은

$$H_n = \{h_1(SPF_{h_1 \cdot n}), h_2(SPF_{h_2 \cdot n}), \dots, h_m(SPF_{h_m \cdot n})\}$$

이 된다.

이렇듯 H_n (가설집합)에 속한 가설들(h_m)은 선택된 가설(h_{MAX})이 검증될 때마다 SPF값이 수정되어진다. 이들중 h_{MAX} 와 연관관계가 없는 가설들은 그 SPF값이 반감된다. 그러므로 h_{MAX} 가 확정된 경우(즉, CF가 매우 큰 경우) 그와 관계가 있는 가설들은 상대적으로 값이 커지고 그러한 영향으로 가설의 선택과 검증이 일련의 깊은 관계를 가진 상황들에 의해 이어지게 된다. 이러한 일관성 있는 가설선택 작용으로 추론작업에 있어 혼돈을 일으키지 않고 효율적으로 행할 수 있다.

이제까지 설명했듯이 특수가능성지수는 가설의 고유한 가능성이며 그에 의해 검증될 가설의 순서가 정해지고 그 수가 음이 되면 가설집합에서 탈락한다.

나) 확실성지수(CF: Certainty Factor)

이는 가설검증후 그 상황중 확실성을 나타내는 지수이다. 이것도 PF(가능성지수)와 마찬가지로 -100~100의 숫자로 되며 가설검증시 다음과 같이 측정된다.

$$C_n = \{S_1(GPF[s_1/h]), S_2(GPF[s_2/h]), \dots, S_n(GPF[s_n/h])\}$$

이라 표시될 수 있으며 $S_1 \dots S_n$ 은 가설 h 의 특성이 되는 상황들이며 $GPF[s_n/h]$ 는 가설 h 에 대한 특성상황 S_n 의 중요도이다. 가설 h 를 검증하여 CF값을 계산하는 수식은 다음과 같다.

$$CF_n = \sum_{k=1}^n (GPF[s_k/h] * CF_{S_k}) / 100$$

여기서 CF_{S_k} 는 S_k , 즉 집합 C_n 중 K 번째 특성의 확실성지수이다(이때 S_k 가 아직 검증되지 않은 상황이라면 그 즉시 검증을 행한다). 또한 $GPF[s_k/h]$ 는 가설 h 에 대한 K 번째 특성의 일반가능성 지수이다.

즉, CF_n (가설 h 의 확실성지수)이란 각 특성들을 검증하여 계산된 CF와 그 특성들의 가설에 대한 GPF값의 곱들의 총합을 의미하며 그에 의해 가설 h 에 대한 확정 여부를 가린다.

```

안녕 하세요
저는 씨. 아이. 에스 인디디

환자의 이름을 입력하십시오 #총 길동
환자의 나이를 입력하십시오 #45
환자의 성별을 입력하십시오 #남성

귀가 병력을 입력하십시오
#당뇨병
#없

환자의 증상을 입력하십시오
#환자가 험기증을 느낌
#환자에 기억력의 장애가 있음
#환자가 피로를 느낌
#환자가 쇠약감을 느낌
#환자의 혈압이 높음
#환자에 두통이 있음
#없

지금부터 추론을 시작합니다

(환자의 질병은 동맥경화임) 을 지금 검증 합니다

환자에 인연이 있습니까? : -100
환자에 혼돈이 있습니까? : 0

(환자의 질병은 뇌졸중임) 을 지금 검증 합니다

환자의 질병은 뇌졸중입니까? : 0

(환자의 질병은 고혈압임) 을 지금 검증 합니다

환자의 목이 검칙됨? : 90
환자의 귀가 울림? : 90

(환자의 질병은 본태성고혈압임) 을 지금 검증 합니다

환자가 호흡이 곤란합니까? : 0
환자가 흥분을 느낌? : 100
환자가 불안함을 느낌? : 100
    
```

중 확

```

모든 추론이 끝났습니다

환자 이름      총 길동
나이          중년
성별          남성

환자의 증상

(환자가 험기증을 느낌) # 100
(환자에 기억력의 장애가 있음) # 100
(환자가 피로를 느낌) # 100
(환자가 쇠약감을 느낌) # 100
(환자의 혈압이 높음) # 100
(환자에 두통이 있음) # 100
(환자의 목이 검칙됨) # 90
(환자의 귀가 울림) # 90
(환자가 흥분을 느낌) # 100
(환자가 불안함을 느낌) # 100
(환자에 인연의 장애가 있음) # 90
(환자의 인자에 혼란이 있음) # 90
(환자가 혈압의 변동이 검칙됨) # 85

환자의 질병

(환자의 질병은 당뇨병임) # 100
결과 (환자의 질병은 고혈압임) # 98

(환자의 질병은 고혈압임) # 98
사본 (환자의 질병은 본태성고혈압임) # 89
현인 (환자의 질병은 당뇨병임) # 100

(환자의 질병은 본태성고혈압임) # 88
소속 (환자의 질병은 고혈압임) # 98
사본 (환자의 질병은 의심본태성고혈압임) # 75

(환자의 질병은 의심본태성고혈압임) # 75
소속 (환자의 질병은 본태성고혈압임) # 88

다음의 상황들은 아직 NIEH지는
있었지만 앞으로 NIEH가 가능성이 있으니 참고 바랍니다.

(환자의 질병은 당뇨병성말초병)
(환자의 질병은 심장부전)
(환자의 질병은 뇌졸중)
(환자의 질병은 뇌혈관)
(환자의 질병은 뇌졸중)
(환자의 질병은 협심증)
    
```

그림 5. 시험구동결과

Fig. 5. Results of imple menting the system.

IV. 시험 구동

본 시스템은 시험구동에 있어서 고철압과 그 합병증에 대한 49개의 질병과 100여개의 증상을 상황으로 하는 지식베이스를 사용하고 있다. 우리는 지식베이스를 구성하기 위해 관계문헌을 이용했는데 지식을 입력하는 과정에서는 거의 어려움이 없었으며 그 추론과 정도 만족할만한 결과를 얻었다.

한편, 본 시스템은 APPLE computer의 APPLE-LISP을 사용하여 수행되었고 지식습득 program¹⁾에 의해 지식을 입력하고 컴퓨터의 보조기억장치에 기억시켜 두었다. 그림 5에 시험구동 예를 보인다.

V. 結 論

여기까지 이번에 제안된 CIS의 지식표현, 그리고 추론과정 등의 설명과 그에 의해 구성된 시스템의 시험구동 결과를 보였다. 시험구동을 수행하면서 나타난 난점이 지식의 복잡성이다. CIS의 각 상황에 대한 지식표현에서는 중복적 표현이 불가피하다. 예를 들어 A라는 상황의 원인이 B라는 상황일때 B의 결과가 A임은 보통의 경우에서는 당연하다. 그러나 이러한 것들 모두를 표현해주어야 함이 번거롭다. 그리고 상황들 각각의 지식들이 rule-based system에서의 rule과 같이 완벽한 독립성을 가지지 못해 지식의 습득, 수정, 추가가 그러한 시스템들에 비해서는 용이하지를 못하다.

그러나 이러한 지식의 복잡성은 본 시스템에서의 기본 개념이 지식의 복잡한 연관관계 표현에 있었기 때문에 어쩔 수 없는 것이며 연관관계를 추론에 이용한 frame-based system들에서 보다는 좀더 유연하고 간결한 지식표현과 추론이 이루어졌다고 보여진다. 이러한 문제는 강력한 지식처리 시스템의 개발에 의해 해소될 수 있으리라 믿는다.

본 시스템의 지식표현은 인간들의 강력히 연결된 지식체계를 되도록 충실히 컴퓨터에 표현하고자 하였다. 이렇듯 강력한 지식체계를 가져야만 인간의 복잡한 지식을 설명할 수 있으며 추론도 용이하게 이루어진다.

실제로 본 시스템은 가설의 설정과 선택이 매우 간단히 수행되었고 추론의 정당성을 위해 꼭 필요한 소속추적과 원인규명이 가능해졌으며 미래에 일어날 수 있는 상황을 예상할 수도 있었다.

부가하여 술어와 한정어에 의한 상황표현은 인간 언어에서의 한정관계와 상통한다고 보기 때문에 어느 정도 수식어의 제한이 따른다면 매우 쉽게 Man-Machine Interface가 구성될 수 있을 것이다. 이도 시험구동에서 실현되었다.¹¹⁾

비록 본 시스템이 전문가의 도움이 없었으며 그 지식베이스도 매우 미흡하여 앞으로 시스템이 확대되었

을 때의 새로운 문제점에 대한 예상을 할 수는 없지만, 본 논문에서 제안된 지식표현과 추론기법이 여러 장점들을 가지고 있기 때문에 여러 분야의 훌륭한 전문가 시스템으로 발전될 수 있으리라 보인다.

參 考 文 獻

- [1] Laurent Sikossy, *Let's Talk Lisp*, Prentice-Hall, 1976.
- [2] A. Barr, E.A. Feigenbaum, *The Handbook of Artificial Intelligence*, William Kaufmann Inc., vol. 1, pp.143-222, vol.2, pp. 79-101, pp.177-222, 1982.
- [3] R. Davis, B. Buchanan, E. Shottliffe, "Production Rules as a Representation for a Knowledge-Based Consultation Program," *Artificial Intelligence*, 8, 1977.
- [4] S.M. Weiss, C.A. Kulikowski, S. Amarel, A. Safir, "A Model-Based Method for Computer-Aided Medical Decision-making," *Artificial Intelligence*, 11, 1978.
- [5] 開原 成允, 小山 照夫, 木村 通男, "専門家の相談に應じるエキスパート・システム," *Nikkei Electronics*, 9, 1982.
- [6] E. A. Feigenbaum, W. J. Clancey, "知識工學-その方向と目標," *數理科學*, no. 214, 4, 1981.
- [7] 최종수, "Knowledge-Based System," 대한전자공학회잡지, vol. 12, no. 2, 4, 1985.
- [8] 최종수, "응용 인공지능(지식공학)," 중앙대학교 대학원 연구논집, 제 4 집, 5, 1985.
- [9] 박중훈, 최종수, "Expert System 구성을 위한 범용 추론 System(CIS)에 관한 연구," 대한전자공학회 하계 종합학술대회 논문집, vol. 8, no. 1, pp. 244~247, 6, 1985.
- [10] 박중훈, 최종수, "상황해석 시스템(CIS)의 인과 및 종속계통에 관한 연구," 대한전자공학회 추계 종합학술대회 논문집, vol. 8, no.2, pp. 626~628, 11, 1985.
- [11] 권오신, 박중훈, 최종수, "상황해석 시스템(CIS)의 상황표현과 한글처리 Interface에 관한 연구," 대한전자공학회 추계 종합학술대회 논문집, vol. 8, no. 2, pp. 720~723, 11, 1985. *

감사의 글

본 연구를 행하는 데 있어 토론에 협조하여 준 화상 정보처리연구실의 인공지능팀에, 특히 끝까지 논문작성상 힘써주어 준 권오신군에 감사한다.