

# 퍼지 프러덕션시스템의 추론방법에 관한 연구

송수섭\*

## A Research on Inference Method in Fuzzy Production System

Soo Sup Song\*

### 요 약

전문가의 지식을 지식베이스화하여 의사결정지원시스템으로 사용하려는 노력이 증대하고 있다. 특히 투자 의사결정과 같은 원인결과의 관계를 명확히 규정할 수 없는 복잡한 영역에서 전문가의 지식베이스는 비전문가의 의사결정에 중요한 조언을 제공할 수 있다. 불확실한 지식을 지식베이스화하는 한 방법으로 퍼지프러덕션시스템이 널리 사용되고 있다. 주식시장과 같은 동태적인 시스템에서 어떤 정보의 중요성은 상황에 따라 변화하는데 이를 정태적인 프러덕션시스템의 규칙으로 지식베이스화하는 것은 불가능하다. 그러나 추론을 수행하는 과정에서 수행당시 각 정보의 중요도에 부응하는 가중치를 부여하여 평가함으로써 정태적인 지식베이스에 동태적인 실제시스템의 특성을 반영할 수 있다. 이는 가중치가 높은 정보에 해당하는 조건명제의 충족정도가 해당규칙의 전체 평가결과에 더욱 중요하게 반영되게하여 좀더 현실성 있는 추론 결과를 얻게한다.

AHP(Analytic Hierachy Process)방법에 의하여 얻어진 정보의 상대적 중요도에 따른 가중치(w)를 해당 정보와 조건명제의 합치정도(Degree of Match : DM)에  $(DM)^w$ 의 형식으로 적용함으로써 퍼지프러덕션시스템에서 정보의 중요도를 반영하여 프러덕션규칙을 평가하는 방법을 제시한다.

주제어 : 퍼지전문가시스템, 퍼지추론, 불확실성

\* 세종대학교 경영대학 정보처리학과 부교수

## I. 서론

조직의 경영에 있어서 의사결정의 문제는 중요한 문제로서 이를 해결하기 위하여 수학적론에 바탕을 둔 각종 계량모형(Quantitative Model)들이 개발되었으며 이러한 노력은 계속되고 있다. 이러한 계량적 모형들의 단점은 입력되는 데이터 및 모형자체의 수학적인 정확성을 필요로 하는 것이다. 그러나 인간의 의사결정은 수치적으로 정확하게 계량화할 수 없는 많은 요소들을 내포하고 있으며, 인간이 어떤 현상을 인지하고 이를 분석하여 미래의 행동방향을 결정하는 것은 수치로 인식하기보다는 과거의 경험이나 현상에 비추어 불분명하게 인식하고 이를 결합하여 의사결정을 하는 경우가 대부분이다.

컴퓨터의 발명과 사용의 확산에 따라 이를 효과적인 의사결정에 활용하려고 하는 노력으로 의사결정지원시스템(Decision Support System : DSS)의 개념이 등장하여 많은 분야에서 활용되고 있으며 활발한 연구가 계속되고 있다. DSS는 경영의사결정에 있어 컴퓨터의 역할에 대한 하나의 관점으로 (1)반구조적(Semistructured)문제에 대한 의사결정과정을 보조하며, (2)의사결정자의 판단을 대치하는 것이 아니라 지원(Support)하는 것이며, (3)의사결정의 효율성(Efficiency)이 아니라 효과성(Effectiveness)을 제고하는 목적을 가지고 있다(Keen and Scott Morton, 1978). 즉 DSS는 경영자가 항상 직면하는 비교적 비구조적인 복잡한 문제에 대한 의사결정에 있어 인간의 능력한계를 보완해 준다. 그리고 최종적인 대안의 선택은 의사결정자의 판단에 의존한다.

근래의 인공지능(Artificial Intelligence)과 전문가 시스템(Expert System) 또는 지식기반시스템(Knowledge-Based System)에 대한 연구는 의사결정에 컴퓨터를 활용하는 새로운 관점을 제공하고 있으며

전문가의 경험, 지식, 의사결정에 이르는 논리적인 과정을 컴퓨터 프로그래밍화함으로써 의사결정의 효과성뿐 아니라 효율성까지 제고시킬 수 있는 능력을 약속하고 있다.

지식기반시스템의 핵심인 프러덕션시스템(Production system)은 해당 문제영역의 지식을 표현하는 주된 형식으로 “if left-hand-side proposition is true, then right-hand-side proposition is true”의 형식으로 이루어진 프러덕션규칙(Production rule)들을 나타내는 프로그램들의 집합이다. 이러한 프로그램들은 두가지 방법으로 작동된다. 규칙의 좌측명제(전제명제(antecedent), 조건명제(condition))가 진실일 경우 우측명제(결과명제(consequent))가 진실 이라고 결론을 내리는 “modus ponens”형식과 도달하려는 목표로부터 이에 가장 근접한 우측명제를 도출하는 좌측명제들을 찾아내려하는 “modus tollens”형식이다.

프러덕션시스템의 가장 중요한 특성중의 하나는 개별적인 IF-THEN 규칙이 전통적인 프로그래밍의 통제구조에 의해서가 아니라 데이터베이스의 형식으로 상호작용 한다는 것이다. 프러덕션시스템은 프러덕션규칙이라는 단편적인 모듈 형태로 지식을 표현하므로 전문가의 지식을 단편적형태(chunk)로 도출할 수 있게 한다. 이렇게 지식을 단편적인 형태의 규칙으로 표현하고 이들을 데이터베이스 형식에 의하여 상호작용하게하는 것은 전문가의 특정임무에 대한 실제 의사결정과정에 가장 근사한 프로그래밍 방법으로 인간의 인식과정을 모형화하는 가장 자연스러운 방법이다(Newell and Simon, 1972).

실세계에서 인간의 지식은 많은 부분 불확실성과 부정확성을 내포하며 또 사람들은 불확실성을 내포하는 정보를 가지고 의사결정을 한다. 따라서 프러덕션규칙에 의하여 인간의 지식을 표현하고 처리하는데는 다양한 불확실성이 존재한다. 이러한 불확실

성을 표현하고 처리하는 하나의 방법이 퍼지셋이론(Fuzzy Set Theory)에 의한 언어변수(Linguistic Variable)를 이용한 시스템이다. 언어변수는 인간이 지식을 표현하는데 가장 자연스러운 수단인 자연어를 코드화하고 수리적 규칙에 맞게 처리할 수 있는 장점을 가지고 있어 Zadeh(1975)가 제창한 이후 다양한 분야에서 연구되고 적용되어 왔다. 전문가의 지식을 언어변수로 표현하고 처리하는 퍼지프러덕션시스템(Fuzzy Production System)에서 여러가지 추론방법이 제창되었으며 이들은 적용영역의 특성에 따라 조금씩 상이한 면을 보여준다(Whalen and Schott, 1983). 이 논문의 목적은 퍼지프러덕션시스템에서 추론을 수행함에 있어 규칙에 포함된 속성들이 결론에 미치는 영향의 상대적 중요도(importance)를 고려하여 규칙들을 평가할 수 있는 하나의 방법론을 제시하는데 있다.

## II. 퍼지셋에 의한 언어변수의 표현

### 2.1 지식기반시스템의 불확실성 처리

인간이 의사결정을 함에 있어 사용할 수 있는 정보는 대부분의 경우 부분적이거나 완전히 확신할 수 없으며, 정보를 표현하는 언어상의 부정확성, 다중원천으로부터 얻어지는 정보들의 상충성 등 완전하지 못하고 결함이 있다. 이러한 불확실성은 실제세계에서 어떤 사상의 발생, 다양한 사상간의 인과관계 또는 상관관계에 관련된 지식에서 공통적으로 존재한다. 즉 인간의 지식은 많은 부분 불확실하며, 또 불확실하고 모호한 규칙과 사실을 내포하고 있는 환경에서 의사결정이 이루어진다. 따라서 의사결정을 지원하기 위한 지식기반시스템에서 부정확하고 불확

실한 인간의 지식을 표현하고 처리하는 수단을 제공해야 한다. 지식베이스에서 지식은 하나의 결과명제(consequent)와 하나 또는 여러개의 전제명제(antecedent)로 구성되는 규칙(Rule)으로 표현되는데 이 규칙에서 지식의 불확실성은 여러가지 요소에 의해 좌우된다.

첫째, 실제현상을 전제명제로 번역하는 과정에서의 불확실성을 들수 있다. 예를 들면 증권투자에 있어서 거래량이 증권시장에 어떻게 영향을 미치는가를 표현하는 규칙이 있다고 할때 거래량은 특정 숫치로 나타나는데 전문가시스템에서는 해당 투자분석 전문가가 이를 주식가격의 등락이라는 결과명제와 연계하여 인식하는 내용으로 표현되어야 한다. 즉 "특정주식의 거래량이 많으면 주식가격이 상승한다"라고 하는 규칙에서 거래량이라는 특정 숫치가 주식가격 상승에 어떤 의미를 갖는가로 1차적으로 해석하여야 한다. 여기서 주식가격의 상승을 가져올 정도로 많은가에 대한 숫치의 정확한 경계가 있을 수 없으며 그렇다고 주식거래량과 가격상승간에 회귀분석모형에 의하여 추정될 수 있는 정확한 숫치적인 인과관계가 존재하는 것도 아니다. 따라서 전제명제 자체에 불확실성을 내포하고 있다.그리고 실제 거래량을 지식기반시스템의 규칙으로 해석하는데 어디서 구분할 것인가 하는 "경계의 문제"가 있다.

둘째, 하나의 규칙에는 두개 이상의 전제명제가 있을 수 있는데 불확실성을 내포한 전제명제들을 결합하여 결과명제와 관련된 전체적인 불확실성을 도출하는데 따른 문제이다. 예를 들면 "특정주식의 거래량이 많고 가격의 상승이 지속적일 경우(상승세의 한계에 도달했을 가능성이 높아) 주식을 매도하는 것이 바람직하다"라는 규칙이 존재할 경우 거래량이 많은 정도의 불확실성과 가격상승의 지속정도에 대한 불확실성을 그리고 두 전제명제의 상호작용을 결합하여 결과명제에 연계된 전체적인 전제명제의 의미로

해석하는데 따른 불확실성을 내포하고 있다. 그리고 의사결정과정에서 같은 결과명제를 갖는 2개 이상의 규칙이 동시에 사용될 때 각 규칙의 불확실성을 결합하여 전체추론의 불확실성을 산출하는 경우에도 같은 결합에 따른 불확실성이 존재한다.

셋째, 전제명제와 결과명제의 원인결과의 관계를 의미하는 추론에 있어서의 불확실성이다. 즉 전제명제의 모든 내용이 정확하다고 할 때 어느 정도의 확신을 가지고 결과명제의 결론을 내릴 수 있는가 하는 것이다. 이의 대표적인 예가 MYCIN 으로 어떤 사실이나 규칙에 대하여 전제명제가 진실이라고 할 때 결과명제의 진실성의 정도를 표현하는  $[-1, 1]$  사이의 값을 부여하여 추론결과의 불확실성 정도를 인식할 수 있게한 것이다.

오랫동안 베이지안 모형(Bayesian model)이 불확실성을 표현하고 이를 포함하는 추론을 수행하는 수리적 방법이었다. 최근 10여년 동안에 확률적 방법과 상이하다고 주장하는 몇가지의 수리적 방법이 제창되었다. 이들 중 Shafer(1976)의 증거이론(evidence theory)과 Zadeh(1978)의 가능성이론(possibility theory)이 중요한 것 들이다. 가능성이론은 Zadeh(1965)가 주창한 퍼지셀이론(fuzzy set theory)에 그 기초를 두고있다.

퍼지셀이론은 전문가의 지식을 규칙으로 표현함에 있어 규칙의 전제명제와 입력내용의 부분적인 합치에 따른 불확실성, 그리고 전제명제의 내용에 따른 결과명제와 인과관계의 불확실성 등 상이한 관점의 불확실성을 수리적으로 표현하여 처리하는 수단을 제공한다. 그리고 퍼지셀이론의 큰 장점의 하나는 퍼지언어변수의 개념을 제공하여 언어적으로 표현되는 지식을 수리적으로 표현하여 처리할 수 있게 한다는데 있다. 사실 퍼지언어변수는 전문가들이 관측된 현상을 전문지식을 기초로 해석하여 관측된 현상이 목적하는 결과에 미치는 영향을 자연어로

표현하는 내용을 수리적으로 표현하여 처리하는데 가장 적절한 도구로 간주되고 있다. 따라서 퍼지언어변수를 도입함으로써 복잡한 규칙들이 간소화 되며 연속량을 갖는 변수를 이산화 하는데서 발생하는 “경계의 문제(Boundary Problem)”를 해결할 수 있다.

## 2.2 언어변수의 표현

언어변수(linguistic variable)의 사용은 Zadeh(1975)가 언어변수에 대한 개념을 발표한 이후 그 관심이 계속 증대되어 왔다. 언어변수는 어학적인 구문법과 의미를 갖는 퍼지변수의 하나이다. 일반적으로 인간의 지식은 수리변수로 나타낼 수 있을 만큼 정확하지 못하며 그렇다고 아무것도 없는 것처럼 모호한 것도 아니다. 이러한 인간의 지식을 컴퓨터에서 표현하여 처리하는데 언어변수의 개념이 대단히 유용하다. 이를 위해서는 언어변수로 표현된 지식을 처리하는 방법이 필요한데 그 전형적인 방법이 “근사추론(Approximate Reasoning)”이다.

퍼지멤버십으로 표시되는 언어변수는 불확실한 사실을 계량화 또는 서술하는 방법을 제공하는 한편 근사추론은 불확실성을 내포하는 If-Then 규칙으로 구성된 지식기반시스템에서 추론을 수행하는 방법을 제공한다. Wenstop(1976)을 시작으로 많은 학자들이 특정 응용분야를 지향하는 특정목적 언어처리시스템을 주창하였으며 일반목적용 시스템으로는 Wenstop's system(1980), FUZZY and L-FUZZY(Freksa, 1982), PRUF(Zadeh, 1981), FRIL(Baldwin & Zhou, 1982) 등이 있다.

이들 언어처리시스템 가운데 Wenstop의 언어처리시스템은 다양한 컴퓨터에서 수행 될 수 있는 APL 코드가 공표 되었을뿐만 아니라 언어적근사 및 근사추론을 위한 강력한 기능을 보유하며 변형과 확

장이 용이하고 융통성 있는 단어들과 연산자들을 제공하고 있어 널리 사용되고 있다. Wenstop의 언어처리시스템은 5가지 기본용어(primary terms)로 "low", "medium", "high", "unknown" 및 "undefined"를 사용하고 있으며 "above"등 14가지의 수식어(hedges), "and"등 6가지의 접속어, "increasingly"등 3가지 추세 표현어, "falling"등 추세의 방향을 나타내는 단어 등을 사용하고 있다.

기본용어들은 멤버쉽함수(membership function)와 그들의 명칭(예를 들면 "high", "low" 등)에 의해서 특징지워진다. 또한 멤버쉽함수는 베이스값(base values)과 이들 베이스값들에 대한 진실값(truth value)으로 표현된다. 예를 들면 증권시장의 고객위탁금이 "많다"는 것을 언어변수로 표현하려면 증권분석가가 생각하는 하한과 상한을 정한 다음 이 범위를 적당한 수의 구간으로 나누고 "많다"라는 개념을 각 구간값에 대한 0 과 1 사이의 진실값을 부여하면 이것이 멤버쉽(membership grade)이고 모든 구간값의 멤버쉽 즉 진실값들의 벡터를 멤버쉽함수라고 한다.

베이스값 5 7 9 11 13 15 17 19 21 23 25(천억)  
진 실 값 0 0 0 0 0 0 0.1 0.3 0.7 1 1

여기서 고객위탁금이 1조 9천억일 경우 많다는 범주에 속하는 정도가 0.3이라는 의미로 해석할 수 있다.

Wenstop 시스템에서는 홀수개의 베이스값을 필요로 하는데 특히 11개의 베이스값을 사용하는 시스템은 11개의 베이스값이 0 에서 100% 까지를 10% 단위로 구간을 정하는 것과 일치하는 유용성이 있어 가장널리 사용되어왔다. 따라서 이 시스템은 표현하려는 영역 또는 정보항목에 관계없이 공통적으로 사용할 수 있는 장점을 가지고 있다. 즉 "많다" 또는 "높다"를

구간값 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0  
멤버쉽 0 0 0 0 0 0 0.1 0.3 0.7 1 1

와 같이 표현하면 이는 거래량이 "많다", PER 가 "높다" 등으로 실제 베이스값에 관계없이 공통적으로 사용될 수 있는 것이다. 이때 해당항목에 대한 베이스값은 전문가의 머리속에 있을 수 있으며 항목별로 다른 베이스값을 11개의 수(0.0 에서 1.0까지)에 대응시키면 된다. 그리고 이 시스템은 상대적으로 표현되는 항목을 표현하는데 유용하다. 예를 들면 특정주식의 주가 상승폭이 "크다"는 표현이 필요한 경우 절대상승가를 베이스값으로 사용하는 것 보다 상승제한폭에 대한 상승액의 퍼센트를 베이스 값으로 하여 표현하는 것이 더 효과적일 수 있다.

Wenstop의 시스템에 있어서 5개의 기본용어를 벡터로 표현하면 다음과 같다.

언어변수 명칭	벡 터 표 현										
LOW :	1	1	0.7	0.3	0.1	0	0	0	0	0	0
MEDIUM :	0	0	0	0	0.4	1	0.4	0	0	0	0
HIGH :	0	0	0	0	0	0	0.1	0.3	0.7	1	1
UNKNOWN :	1	1	1	1	1	1	1	1	1	1	1
UNDEFINED :	0	0	0	0	0	0	0	0	0	0	0

이러한 언어변수의 의미는 베이스값에 대응하는 숫치의 벡터로 표현되며 각 숫치들이 의미하는 것은 이에 대응하는 베이스값에 대한 가능성(possibility)을 나타낸다. 원래 멤버쉽함수는 연속함수이며 베이스값은 이 연속값을 근사적으로 표현하기 위하여 사용되는 것이다. 따라서 표현하려는 지식분야 및 해당 속성에 따라 베이스값의 범위 및 근사를 위한 이산값의 수가 달라질 수 있다.

Gracia *et al.*(1992)는 부정확한 데이터라 가지고 어떤 시스템의 실패위험을 평가하는 방법을 제시하

면서 6개의 요소를 갖는 벡터로 언어변수를 표현하였는데 기본용어의 멤버십은 아래와 같으며 이 기본용어에 부가하여 “very, pretty, sort of, rather, slightly, low to medium, medium to high” 등의 수식어를 사용하고 있다.

high=(0/1, 0.2/2, 0.4/3, 0.6/4, 0.8/5, 1/6).

medium=(0.1/1, 0.3/2, 1/3, 1/4, 0.3/5, 0.1/6)

low=(1/1, 0.8/2, 0.6/3, 0.4/4, 0.2/5, 0/6)

Gracia *et al.*의 언어처리시스템은 Wenstop 시스템에 비하여 언어변수들의 멤버십값의 중복이 커서 용어들간의 경계가 더욱 불명확한 것을 보여준다. 어떤 시스템이 더 바람직할 것인가는 적용 업무영역에 따라 다를 것이다. 즉 언어변수의 표현을 과도하게 세분화하는 것은 인간의 서로 다른 표현간의 구분능력의 한계를 초과하여 실용성이 없으며 또한 너무 큰 범주로 구분하며는 인간이 인식할 수 있는 구분범주를 표현할 수 없어 비효율적인 시스템이 될 수 있다. 그러므로 적용 업무영역에 따라 적절한 배이스값과 멤버십값을 찾아내는 것은 많은 노력이 필요하고 이는 퍼지언어변수를 이용한 지식기반시스템 구축의 중요한 관건이 된다고 할 수 있다.

### III. 퍼지 프러덕션시스템(Fuzzy Production System)

프러덕션시스템은 프러덕션규칙(production rule)이라고 하는 IF(condition)-THEN(action) 형식을 주로 사용하는 컴퓨터 프로그램이다. 이러한 프러덕션시스템을 인간의 문제해결 행태를 묘사하는데 사용하려는 관심은 Newell & Simon(1972)이 인간의

지식은 거의 독립적으로 작용하는 “논리적 집단(logical chunk)”으로 조직화 된다는 연구발표가 있음으로부터 이다. 하나의 프러덕션규칙에서 IF문은 어떤 사람이 특정의 지식집단(knowledge chunk)을 회생(retrieve)시키는데 필요한 조건을 의미하며 THEN문은 이 회생된 지식을 적용한 결과를 의미한다. 이러한 프러덕션시스템은 다양한 분야에서 전문가의 지식을 도출하고 표현하여 지식베이스화 하는데 널리 사용되고 있다.

일반적인 프러덕션시스템은 아리스토텔리안(Aristotelian)논리에 근거하여 구성되었기 때문에 어떤 규칙에서 IF문의 조건이 일치하는 경우에 “all or nothing”의 형식으로 그 규칙이 작동된다. 따라서 모든 변수는 명목척도(nominal scale)로만 처리되며 숫자척도로 측정된 변수를 이러한 프러덕션시스템에서 사용하기 위해서는 숫치를 구간화하고 규칙에서는 측정된 데이터가 이 구간에 속하는가 아닌가를 검사하여 규칙을 작동할 것인지 아닌지를 결정하게 된다.

이러한 “all or nothing” 형식으로부터 야기되는 문제를 경감시키기 위하여 MYCIN과 같은 성공적인 프러덕션시스템에서는 다양한 형태의 불확실성 또는 근사(approximation)를 묘사하기 위한 방법을 도입하였다. 그러나 이러한 시스템의 수정은 프러덕션시스템의 모듈적(Modularity) 특성에 따른 일부 장점들을 훼손하는 원인이 되고있다. 퍼지프러덕션시스템은 프러덕션시스템의 장점을 최대화 하면서 “all or nothing”의 문제를 해결하는 방법으로 대두되었다.

퍼지프러덕션시스템에서는 IF-THEN 관계를 수행하기 위하여 아리스토텔리안 논리가 아닌 다수값논리(many valued logic)를 사용한다(Mizumoto 1982 ; Prade 1983 ; Lesmo *et al.* 1985 ; Cho *et al.* 1992). 즉 IF문의 조건의 일치성을 평가함에 있어서 완전일치와 완전불일치 사이의 일치정도를 인식하여 평

가한다. 입력데이터와 IF문의 조건간에 일치정도에 따라 적절히 약화된 THEN문의 결과를 수행한다. 이러한 작동이 이루어 질 수 있도록 퍼지프러덕션 규칙에 있어서 전제와 결과는 가능한 값들에 대한 퍼지셀의 구성인자로 표시되는 퍼지명제로 이루어진다. 퍼지명제의 간단한 예를 들면 “John is tall”이라는 명제는 John의 키가 180Cm 일 경우 진실에 가까우나 John의 키가 200Cm 일 경우는 이 보다 그 진실의 정도가 훨씬 높다.

전통적인 규칙기반시스템(Rule Based System)에 있어서 프러덕션규칙은 데이터가 규칙의 전제조건에 완전히 일치하지 않는 경우에 그 규칙은 추론에 전혀 영향을 미치지 않는다. 또한 규칙들은 순차적으로 작동되며 두가지 이상 규칙의 전제조건이 동시에 만족할 경우 어느 한쪽을 먼저 작동 시키도록 하는 충돌해결 방법이 필요하다. 이에 반해서 퍼지프러덕션시스템에 있어서는 한번 작동에서 모든 규칙이 수행되며 데이터가 규칙의 퍼지조건과 일치정도가 완전 불일치로부터 완전 일치까지의 범위로 나타난다. 전제조건이 완전 만족될 경우 규칙작동의 결과는 결과명제와 정확하게 일치하며 부분적으로만 만족할 경우에는 작동결과는 조건만족의 불완전성을 반영하여 규칙의 결과명제와 유사하나 좀더 확실치 않은 상태로 나타난다. 그리고 전제조건이 전연 만족되지 못할 경우 작동결과는 아무것도 없는 즉 결과명제의 변수에 아무런 제약을 초래하지 않는다.

퍼지프러덕션규칙에 의하여 지식을 표현하는 것은 언어변수를 도입함으로써 더욱 편리하게 되었다. 언어변수는 어떤 수준의 불확실성을 내포한 지식도 가장 근사하게 표현하여 처리할 수 있는 융통성이 있어 지식의 표현에 언어변수를 사용함으로써 특정 분야 전문가로부터 집단화된 정보를 도출하는 것이 더욱 용이해 졌다. 특히 경영분야와 같이 경험적인 지식이 많은 부분을 차지하는 전문가의 지식을 도

출하고 표현하는데 유용하다. 프러덕션규칙을 퍼지셀으로 표현하는 장점은 Buchanan(1983)이 논의한 바와 같이 여러 규칙들을 중복하여 적용할 수 있는데 기인한다. 퍼지셀으로 표현된 규칙들에서 두가지 이상 규칙의 경계선상의 데이터에 대하여 강제로 어느 한 규칙의 조건에 맞는 것으로 간주하지 않고 그 정도를 반영하여 명시적으로 처리할 수 있다. 결과적으로 각 규칙은 적어도 부분적으로 여러가지의 상황에 대처할 수 있어 지식베이스를 단순화 할 수 있다.

일단 퍼지규칙기반시스템이 작동되면 어떤수준의 불확실성도 처리할 수 있는 언어변수의 기능은 시스템에 데이터를 입력하는데 있어 유용한 융통성을 제공한다. 경영관리 분야의 예를 들면 금전거래 대금은 아주 정확하게 표현되는 반면에 경쟁우위의 정도 또는 융통성 등은 덜 정확하게 표현된다. 경영 의사결정에 있어서도 단기예산은 정확하게 표현되는 반면에 경영정책과 같은 광범위한 내용은 정확하게 측정될 수 없는 형태로 표현된다. 퍼지규칙기반시스템은 입력데이터의 불확실성을 인식하고 합리적으로 처리하여 이에 상응하는 적절한 결론의 신뢰성을 도출하므로 이러한 상황에 특히 유용하다.

프러덕션규칙을 평가함에 있어 입력자료의 중요도는 조건명제가 결과명제에 미치는 영향의 정도로 정의할 수 있다. 예를 들면 “이자율이 하락하고 고객위탁금이 증가하면 주가가 상승한다”는 규칙을 가정할 때 주가상승의 정도는 이자율의 하락정도와 고객위탁금의 증가정도에 따라 좌우되며 주가상승이라는 결과명제의 관점에서 고객위탁금의 증가와 이자율의 하락가운데 더욱 강하게 주가상승에 영향을 미치는 것이 더 중요한 인자라고 할 수 있다. 이러한 인자의 중요도는 상황에 따라 변하는 역동적인 특성을 갖는다. 즉 상황에 따라 이자율의 하락이 고객위탁금의 증가보다 더 중요한 의미를 가질수 있거나 또는 그 반대가 될 수 있다. 그러나 수시로 변화되는

조건명제의 중요도는 정태적인 규칙기반 지식베이스의 규칙으로 저장할 수는 없다. 그러므로 결과를 얻기위한 지식베이스의 평가시 동태적인 인자의 중요도를 반영하여 평가하는 수단을 제공하는 것이 바람직하다.

#### IV. 퍼지 프리덕션규칙의 평가

의사결정은 하나의 분류문제로 생각할 수 있다. 즉 미래에 일어날 수 있는 결과를 여러가지 부류(class)로 분류하고 현재 주어진 정보를 해석하고 그 결과를 결합하여 미래에 어떤 부류의 현상이 발생할 것인가를 예측하여 이를 바탕으로 의사결정을 하는 형태를 취하고 있기 때문이다. 투자자의사결정의 예를 들면 가용한 정보를 해석하여 앞으로 해당 주식의 가격이 상승할 것인가, 포함일 것인가, 또는 하락할 것인가로 분류하여 매수, 매도, 보유등의 의사결정을 하게 된다. 미래 어떤 부류의 현상이 발생할 것인가는 여러가지 상황변수(속성)의 영향을 받으며 특정영역에서 전문가는 이러한 인과관계에 대한 지식을 가지고 있다. 이와 같은 지식은 지식베이스화하여 전문가가 아닌 사람들에게 일관성 있는 조언을 제공하는데 사용될 수 있다.

규칙을 기반으로하는 지식베이스는 다음과 같이 전제명제의 속성값과 그에 따른 결과의 형태로 표현되는 규칙들로 이루어 진다.

IF  $X_1$  is  $A_{11}$  and  $X_2$  is  $A_{12}$  and ..... and  $X_n$  is  $A_{1n}$

THEN Class is  $C_1$ ,

IF  $X_2$  is  $A_{21}$  and  $X_2$  is  $A_{22}$  and ..... and  $X_n$  is  $A_{2n}$

THEN Class is  $C_2$ ,

IF  $X_3$  is  $A_{31}$  and  $X_2$  is  $A_{32}$  and ..... and  $X_n$  is  $A_{3n}$

THEN Class is  $C_3$ ,

IF  $X_1$  is  $A_{m1}$  and  $X_2$  is  $A_{m2}$  and .....and  $X_n$  is  $A_{mn}$

THEN Class is  $C_m$ .

여기서 각규칙은 최대  $n$  개의 조건명제를 가질 수 있으며  $C_1, C_2, C_1, \dots, C_m$ 은 결론을 도출하고자 하는 집단의 부류를 말하는 것으로 증권투자의 경우 주가의 상승, 하락, 포함, 강보합, 약보합, 급상승, 완만상승, 급하락, 완만하락 등의 부류를 들 수 있다. 또  $X_1, X_2, X_3, \dots, X_n$ 은 부류를 규정하는데 영향을 미치는 속성들로 예를 들면 경제성장률, 무역수지, 통화증가율 등 전체경제관련 지표, 산업매출증가율, 산업수출증가율, 산업평균수익률 등 산업관련지표, 개별기업의 PER, 매출증가율, 수출동향 등 개별기업 관련지표들을 의미한다. 즉 이러한 속성들의 관측치 또는 기대치에 따라 해당기업의 주식이가격이 어떤 부류에 속할 것인가를 예측하는 것이다. 따라서  $A_{11}, \dots, A_{mn}$ 은 각 속성들의 관측치 또는 기대치를 의미하며 퍼지프리덕션규칙에서는 일반적으로 “높다/낮다, 많다/적다” 등 퍼지언어변수로 표현된다. 이와 같은 지식베이스를 구축한 후에는 의사결정을 위하여 조언이 필요한 당시의 각 속성들의 관측치 또는 기대치를 수집 또는 추정하여 지식베이스에 입력하여 규칙들을 작동시킴으로써 원하는 조언, 즉 원하는 주식투자관련 예측정보를 획득하여 투자자의사결정에 사용할 수 있다.

이러한 규칙들로 구성된 프리덕션시스템에서 어느 한 규칙의 결과명제(상승, 하락 등)는 “crisp”한 현상으로 규정할 수 있으며 전제명제들(경제성장률, 무역수지 등)은 퍼지셀으로 표현되는 언어변수로 규정할 수 있다. 그리고 결과명제들은 전체집합을 구성하는 배타적 부분집합으로 규정할 수 있다. 따라서 이러한 규칙을 평가함에 있어 결과명제의 불확실성은 전제명제들을 결합하여 평가한 결과의 불확실성으로



나타날수 있다. 즉 첫번째 규칙에서  $C_1$  의 불확실성은 현재 상황을 표현하는  $A_{11}^* \dots A_{1n}^*$  라는 정보를 규칙의 전제명제인  $A_{11} \dots A_{1n}$  과 비교평가하여 전제명제 전체의 불확실성을 산출한 값이 된다. 이 경우 전제명제의 인자들인  $X_1 \dots X_n$  이  $C_1$  에 미치는 중요도를 고려하여 규칙을 평가함으로써 동태적으로 변하는 상황을 좀더 정확하게 분석할 수 있다.

#### 4.1 합치정도(Degree of Match : DM)의 계산

지식베이스에서 규칙을 표현하는 복합 조건명제들의 진리값을 구하고 이를 결합하여 규칙 전체의 신뢰도를 도출하는 여러가지 방법들이 제안되었다 (Lesmo *et al.*(1985), Cayrol *et al.*(1982), Leung and Lam(1988), Cho *et al.*(1992)). 실제 추론을 수행함에 있어서 속성들의 입력자료가 규칙의 조건명제와 완전히 합치하는 경우보다는 부분합치의 경우가 많다. 또한 투자 의사결정의 경우 입력자료들은 거래량, 매출증가율 등과 같은 정확한 수치로 표현되는 것이 있고 투자자들의 투자심리등 모호하게 표현할 수 밖에 없는 데이터가 혼입되어 있다. 따라서 입력자료와 규칙의 조건들과의 부분적 합치를 처리하는 방법이 필요하다. Cho *et al.*(1992)은 규칙의 조건부와 결론부가 정확(precise)하거나 부정확(imprecise)하거나 어떠한 조합에서도 일관성 있게 입력자료와 합치정도를 도출하는데 사용될 수 있는 방법을 제안하였는데 그 계산방법은 다음과 같다.

$$CERT_+ = \text{MAX} \{ \text{MIN}(\mu_A, \mu_{A^*}) \}$$

$$CERT_- = \text{MAX} \{ \text{MIN}(1 - \mu_A, \mu_{A^*}) \}$$

여기서  $\mu_A : X \text{ is } A$  라는 규칙의 조건명제에서  $x$  가  $A$  에 속할 멤버십

$\mu_{A^*} : X \text{ is } A^*$  라는 입력자료에서  $x$  가  $A^*$  에 속할 멤버십

예를 들어 규칙과 이에 대한 수집된 입력자료가 다음과 같다고 하자.

규칙 : IF  $X$  is  $A$  then  $Y$  is  $B$

입력자료 :  $X \text{ is } A^*$

이때 속성  $X$ 의 베이스값이  $X = \{10, 20, 30, 40\}$ 와 같고  $A = \{1/10, 0.6/20, 0.6/30, 0.2/40\}$ ,  $A^* = \{1/10, 0.4/20, 0.7/30, 0.8/40\}$ 와 같은 퍼지변수값을 취한다면 그 결과는 다음과 같다.

$$CERT_+ = \text{MAX}[\text{MIN}(\{(1, 1), (0.6, 0.4), (0.6, 0.7), (0.2, 0.8)\})]$$

$$= \text{MAX}[1, 0.4, 0.6, 0.2] = 1$$

$$CERT_- = \text{MAX}[\text{MIN}(\{(0, 1), (0.4, 0.4), (0.4, 0.7), (0.8, 0.8)\})]$$

$$= \text{MAX}[0, 0.4, 0.4, 0.8] = 0.8$$

합치정도(Degree of Match : DM)는  $CERT_+$  와  $CERT_-$  를 이용하여 다음 세가지 경우로 구분되어 계산된다. Cho *et al.*(1992)에서는  $[+1, -1]$ 의 범위를 취하는 DM과  $[0, 1]$ 의 범위를 취하는 DM 의 두가지 척도를 제시하고 있는데 후자는 Cayrol(1982), Leung (1988)의 방법과 같은 결과를 나타내고 있다. 이 논문에서는 속성정보의 중요도를 규칙의 평가에 적용하기 위하여 후자 즉  $[0, 1]$  범위를 취하는 DM 을 사용한다. 여기서  $DM=1$  은 완전히 합치되는 경우를 의미하며  $DM=0$  은 전혀 합치되지 않는 경우를 그리고  $DM=0.5$  는 중립을 의미한다.

$$CERT_+ \geq 0.5 \text{ 와 } CERT_- \leq 0.5 \text{ 인 경우 } DM = CERT_+$$

$$CERT_+ \geq 0.5 \text{ 와 } CERT_- \geq 0.5 \text{ 인 경우}$$

$$DM = (CERT_+ - CERT_-) + 0.5$$

$CERT_+ \leq 0.5$ 와  $CERT_- \geq 0.5$ 인 경우  $DM = 1 - CERT_-$

상기 예는 두번째의 경우에 해당하므로  $DM = (1 - 0.8) + 0.5 = 0.7$  이라고 계산되며 그 의미는 “IF X is A THEN Y is B”라는 규칙에서 X is A\* 라는 정보를 입력하여 평가할 때 Y is B 라는 결론부의 적합성은 0.7 정도라고 할 수 있다는 것이다.

각 규칙에는 여러개의 조건들이 “AND” 로 연결되어 하나의 결론부를 가질수 있는데 이 경우 해당 규칙의 전체적 DM 은  $\{DM_1, DM_2, DM_3, \dots, DM_n\}$ 으로 계산되며 여러개의 규칙들이 같은 결론부를 가지는 경우 지식베이스 전체에서의 DM 은  $\{DM \text{ of Rule } 1, DM \text{ of Rule } 2, \dots, DM \text{ of Rule } m\}$ 로 계산된다. 따라서 하나의 규칙에 OR 로 연결된 조건명제들을 표현하고자 할 때 퍼지프러덕션시스템에서는 별도의 규칙으로 분리하더라도 모든 규칙이 작동되기 때문에 MAX 연산자에 의해서 결합되는 목적을 달성할 수 있다. 즉 하나의 규칙 내부에서는 “OR” 로 연결되는 조건들을 배제할 수 있다. 예를 들어 “ $X_1 \text{ is } A_1 \text{ OR } X_2 \text{ is } A_2 \text{ THEN Y is B}$ ” 라는 하나의 규칙을 “ $X_1 \text{ is } A_1 \text{ THEN Y is B}$ ” 와 “ $X_2 \text{ is } A_2 \text{ THEN Y is B}$ ” 의 두개의 규칙으로 만들면 매번 두가지 규칙이 작동되기 때문에 “OR” 로 연결된 경우와 같은 결과를 얻는다.

전문가의 지식을 규칙화 하여 지식베이스를 구축함에 있어서 세분화된 규칙을 도출하는 것은 많은 노력이 필요하며 전문가 역시 인간의 제한된 기억 능력 및 정보처리능력으로 인하여 실제 발생하는 모든 현상을 구분하여 규칙화한다는 것은 불가능한 일이다. 따라서 부분적합성에 의한 퍼지프러덕션규칙의 평가는 지식베이스를 대표적인 패턴으로 단순화하더라도 다양한 입력자료(정보)들을 수용하여 평가할 수 있는 장점을 제공한다.

## 4.2 정보의 중요도 평가

상기의 합치정도(DM) 계산에 있어서 입력자료는 지식베이스에 입력하여 원하는 조언을 얻고자 수집한 정보자체 또는 수집한 정보를 해석한 결과를 의미한다. 주식시장의 예를 들면 주식이격은 주식투자자들의 기대심리(expectation)에 따라 달라진다. 따라서 주식시장에서는 어떤정보가 다른 정보보다 더 중요하게 해석되고 주식이격에 크게 영향을 미친다. 즉 상황에 따라서 어떤부류의 정보가 다른부류의 정보에 비교하여 더 중요하게 작용하는 것이다. 이러한 중요도는 상황에 따라 변화되기 때문에 확정적으로 지식베이스화하여 규칙에 반영할 수 없다. 따라서 평가당시 정보들의 중요도를 고려하여 규칙들의 평가에 이를 반영할 수 있도록함으로써 좀더 시의 적절한 조언을 도출할 수 있다.

Ibrahim and Ayyub(1992)는 다중기준(multiple criteria :  $C_i$ )에 의거 대안( $X_i$ )을 평가하고 최적대안을 선택하는 의사결정환경에서 대안들의 각 기준별 충족도를  $[0, 1]$  범위값으로 규정하고 AHP(Analytic Hierach Process : Saaty, 1980, 1987)방법을 사용하여 기준들의 중요도를 나타내는 가중치( $w_i$ )를 구하여 각 대안별로  $\{MAX(\cap(C_i)^{w_i})\}$ 를 계산하여 이값이 제일 큰 대안을 선택함으로써 대안 평가기준들의 중요도를 감안하여 대안들의 우열을 결정하는 방법을 제시하였다. AHP. 방법은 인간의 제한된 정보처리 능력을 극복하기 위하여 매번 두가지 속성간의 정성적인 반복적 비교를 통하여 전체 속성들의 상대적인 정량적 가중치를 구하는데 널리 사용되고 있는 방법이다.

퍼지프러덕규칙에서 정보의 중요도를 규칙의 평가시 반영하기 위하여 Ibrahim and Ayyub(1992)가 제안한 방법을 적용할 수 있다. 즉 조건명제들의 합치정도를 계산하고 한 규칙 전체의 합치정도를 계

산함에 사용되는 MIN 연산자에서 가장 작은 값이 최종결과에 반영되며, 한 규칙에 포함된 조건명제들의 합치정도는  $0 \leq DM_i \leq 1$  범위값을 취하므로  $(DM_i)^w$ 의 값은 w 값이 커질수록 작아지는 특성을 이용하여 각 입력자료들의 중요도를 평가하는 가중치를 구하여 w 값으로 사용함으로써 중요도가 높은 정보가 최종결과에 더욱 강하게 영향을 미칠수 있게 하는 방법이다.

상기의 절차를 예를 들어 설명하기 위하여 다음과 같은 규칙과 입력자료(정보)가 있다고 하자. 여기서 입력자료는 지식베이스의 규칙을 적용하여 평가하려는 대상의 속성( $X_1, X_2, X_3$ )에 대한 값을 의미한다.

규칙 : IF  $X_1$  is  $A_1$  and  $X_2$  is  $A_2$  and  $X_3$  is  $A_3$  then Y is B

입력자료 :  $X_1$  is  $A_1^*$ ,  $X_2$  is  $A_2^*$ ,  $X_3$  is  $A_3^*$

그리고  $X_1, X_2, X_3$ 의 베이스값이 공히 {10, 20, 30, 40}으로 주어지고 각 속성의 멤버십값이

$$A_1 = \{1/10, 0.6/20, 0.6/30, 0.2/40\}$$

$$A_2 = \{0.4/10, 0.8/20, 1/30, 0.6/40\}$$

$$A_3 = \{0.8/10, 1/20, 0.5/30, 0.2/40\}$$

입력자료 1 :  $A_{11}^* = A_{21}^* = A_{31}^*$

$$= \{1/10, 0.8/20, 0.5/30, 0.2/40\}$$

입력자료 2 :  $A_{12}^* = A_{22}^* = A_{32}^*$

$$= \{0.2/10, 0.7/20, 1/30, 0.6/40\}$$

라고 주어졌다고 가정하고 각 정보의 중요도가 같다고 가정할 경우 전체 적합도는 다음과 같이 계산된다.

입력자료 1 :

$$\text{조건 1 : } CERT_+ = 1, CERT_- = 0.4 \Rightarrow DM_{11} = CERT_+ = 1$$

$$\text{조건 2 : } CERT_+ = 0.8, CERT_- = 0.6 \Rightarrow$$

$$DM_{21} = (CERT_+ - CERT_-) + 0.5 = 0.7$$

$$\text{조건 3 : } CERT_+ = 0.8, CERT_- = 0.4 \Rightarrow DM_{31} = CERT_+ = 0.8$$

$$\text{전체적합도} = \min(DM_{11}, DM_{21}, DM_{31}) = 0.7$$

입력자료 2 :

$$\text{조건 1 : } CERT_+ = 0.6, CERT_- = 0.6 \Rightarrow$$

$$DM_{12} = (CERT_+ - CERT_-) + 0.5 = 0.5$$

$$\text{조건 2 : } CERT_+ = 1, CERT_- = 0.4 \Rightarrow DM_{22} = CERT_+ = 1$$

$$\text{조건 3 : } CERT_+ = 0.7, CERT_- = 0.6 \Rightarrow$$

$$DM_{32} = (CERT_+ - CERT_-) + 0.5 = 0.6$$

$$\text{전체적합도} = \min(DM_{12}, DM_{22}, DM_{32}) = 0.5$$

다음 중요도를 고려할 경우는 다음과 같이 계산할 수 있다.

i) 속성비교행렬의 작성(Saaty, 1980, 1987)

이 행렬은  $X_1$ 이  $X_2$ 보다 약간 중요(Weak Importance)하며  $X_3$  보다는 상당히(Strong) 중요함과 매우(Very) 중요함의 중간정도의 중요도를 보이는 것이며  $X_2$ 는  $X_3$  보다 같은정도(Equal)와 약간(Weak) 중요함의 중간정도임을 나타내며 행렬의 하반 삼각부분(Lower Triangle)은 상반부의 역수를 입력한 것이다.

속성	$X_1$	$X_2$	$X_3$
$X_1$	1	3	6
$X_2$	1/3	1	2
$X_3$	1/6	1/2	1

1 : Equal importance

3 : Weak importance of one over another

5 : Strong importance

7 : Very strong importance

9 : Absolute importance

2, 4, 6, 8 : Intermediate values between two adjacent scale judgements

ii) 고유치(Eigen Value)와 고유벡터(Eigen Vector)

를 구하고 제일 큰 고유치에 대응하는 고유벡터를 정규화(Normalize : 벡타요소의 합이 1이 되도록 하는것)하고 여기에다 속성의 수를 곱하여 이를 적합도 값(DM)에 자승하는 가중치( $w_i$ )로 사용한다. 퍼지프 러덕션시스템에서 이러한 가중치를 구하는 것은 전 위치처리를 두어 사용자가 상기 속성비교행렬을 입력하면 계산되어 처리할 수 있게 하면 된다.

최대 고유치 : 3.00  
 대응 고유벡터 : 0.937 0.312 0.156  
 정규화 고유벡터 : 0.667 0.222 0.111  
 가중치 :  $3 * (0.667 \ 0.222 \ 0.111) = 2 \ 0.667 \ 0.333$

입력자료 1의 전체적합도  
 $= \text{MIN}((DM_{11})^2, (DM_{21})^{0.667}, (DM_{31})^{0.333}) = 0.788$   
 입력자료 2의 전체적합도  
 $= \text{MIN}((DM_{12})^2, (DM_{22})^{0.667}, (DM_{32})^{0.333}) = 0.25$

상기 행렬에서 속성  $X_1$ 이 가장 중요하고 입력자료 1에서 이에따른 적합도( $DM_{11}$ )가 1로서 가장커서 전체적인 적합도는 0.7에서 0.788로 증가되었으며 입력자료 2에 대해서는  $X_1$ 의 적합도가 0.5 로서 다른 속성의 적합도 보다 낮아 전체적합도는 0.5에서 0.25로 감소되어  $X_1$ 의 상대적인 중요성이 반영되었다고 할 수 있다. 즉 중요도가 높은 속성의 적합도가 전체적합도에 더 크게 영향을 미친다.

이제 입력자료의 중요도가 변화될 때 적합도가 어떻게 변화하는가를 보기위하여 다음과 같은 속성 비교행렬로 상기 예의 중요도가 변화되었다고 하자. 이 행렬은  $X_2$ 가  $X_1$  보다 상당히(Strong) 중요함과 매우(Very) 중요함의 중간정도 중요하고  $X_3$ 는  $X_1$  보다 동등함(Equal) 과 약간(Weak)의 중간정도 중요함을 나타낸다. 따라서 가장 중요한 속성은  $X_2$ 이다.

이에 대한 계산내용은 다음과 같다.

속성	$X_1$	$X_2$	$X_3$
$X_1$	1	1/6	1/3
$X_2$	6	1	2
$X_3$	3	1/2	1

최대 고유치 : 3.0  
 대응 고유벡터 : 0.1474 0.8847 0.4423  
 정규화 고유벡터 : 0.1 0.6 0.3  
 가중치 :  $3 * (0.1 \ 0.6 \ 0.3) = 0.3 \ 1.8 \ 0.9$

입력자료 1의 전체적합도 :  
 $\text{MIN}((DM_{11})^{0.3}, (DM_{21})^{1.8}, (DM_{31})^{0.9}) = 0.53$   
 입력자료 2의 전체적합도 :  
 $\text{MIN}((DM_{12})^{0.3}, (DM_{22})^{1.8}, (DM_{32})^{0.9}) = 0.63$

입력자료 1의 전체적합도는 0.7 에서 0.53으로 감소 되었으며 입력자료 2의 전체적합도는 0.5에서 0.63으로 증가되어 입력자료 1과 입력자료 2의 전체적합도의 순서가 전도되었다. 이는 두번째 조건명제의 적합도가 입력자료 1의 경우 0.7이며 입력자료 2의 경우 1이고 두번째 속성이 가장 중요하여 적합도가 큰 속성의 중요도가 반영된 것으로 해석할 수 있다.

다음 세번째 속성( $X_3$ )이 가장 중요한 경우의 속성 비교행렬과 그 계산결과는 다음과 같다.

속성	$X_1$	$X_2$	$X_3$
$X_1$	1	1/3	1/6
$X_2$	3	1	1/2
$X_3$	6	2	1

최대 고유치 : 3.0  
 대응 고유벡터 : 0.1474 0.4423 0.8847

정규화 고유벡터 : 0.10 0.30 0.60

가중치 : 0.30 0.90 1.80

입력자료 1의 전체적합도 :

$$\text{MIN}((DM_{11})^{0.3}, (DM_{21})^{0.9}, (DM_{31})^{1.8})=0.67$$

입력자료 2의 전체적합도 :

$$\text{MIN}((DM_{12})^{0.3}, (DM_{22})^{0.9}, (DM_{32})^{1.8})=0.40$$

이 경우 가장 중요한 속성이  $X_3$ 이며 이에 대한 입력자료 1의 적합도(0.8)가 입력자료 2의 적합도(0.6)보다 커서 전체적합도의 차이가 중요도가 동일한 경우보다 증폭 되었다(0.2 → 0.27). 역시 속성의 중요도가 반영된 결과라고 할 수 있다.

좀더 다양한 경우를 검토하기 위하여 한 규칙에서 각 입력자료에 대한 속성들의 적합도를 임의로 가정하고 속성들의 중요도는 상기의 예에서 사용한 세가지 속성행렬의 경우를 적용하여 전체적합도를 계산해 보면 아래와 같다. 입력자료 3, 4, 5의 경우에도 가중치가 가장 큰 속성의 초기적합도의 순서가 가중치 적용후 전체적합도의 순서와 일치하고 있어서 속성의 중요도가 전체적합도에 반영 되었음을 알 수 있다.

속 성	$X_1$	$X_2$	$X_3$	전체적합도
가중치	1	1	1	
입력자료 3 :	0.3	0.4	0.8	0.3
입력자료 4 :	0.6	0.7	0.4	0.4
입력자료 5 :	0.8	0.5	0.2	0.2

속 성	$X_1$	$X_2$	$X_3$	전체적합도
가중치	2	0.6667	0.3333	
입력자료 3 :	0.09	0.63	0.93	0.09
입력자료 4 :	0.36	0.79	0.74	0.36
입력자료 5 :	0.64	0.63	0.58	0.58

속 성	$X_1$	$X_2$	$X_3$	전체적합도
가중치	0.3	1.8	0.9	
입력자료 3 :	0.70	0.29	0.82	0.29
입력자료 4 :	0.86	0.53	0.44	0.44
입력자료 5 :	0.94	0.29	0.23	0.23

속 성	$X_1$	$X_2$	$X_3$	전체적합도
가중치	0.3	0.9	1.8	
입력자료 3 :	0.70	0.54	0.67	0.54
입력자료 4 :	0.86	0.73	0.19	0.19
입력자료 5 :	0.94	0.54	0.06	0.06

AHP에 의한 가중치는 속성들의 상대적인 중요성을 의미하므로 이를 적용하여 계산된 합치정도(DM) 역시 상대적인 성격을 갖는다. 의사결정은 여러가지 의사결정대안 가운데 어떤 대안을 선택할 것인가를 결정하는 경우가 많으므로 상대적인 합치정도의 차이만을 가지고도 의사결정에 도움이되는 조건을 제공할 수 있으므로 상기에 제시한 방법은 특정영역에서 의사결정을 돕기위한 퍼지퍼덕션시스템을 기반으로하는 지식베이스에서 퍼덕션규칙의 평가에 유용하게 사용될 수 있다.

## V. 요약 및 결론

전문가시스템은 특정분야 전문가의 지식과 이를 처리하여 의사결정에 이르는 논리적 과정을 컴퓨터 프로그램화하여 의사결정자를 대치하는 것을 목표로 연구가 진행되어 왔다. 이의 일환으로 인간의 불확실한 지식을 코드화하여 일관성있게 처리하기 위한 퍼지퍼덕션시스템이 제안되었다. 퍼지퍼덕션시스템은 인간이 관찰한 현상이나 수집된 정보를 해

석하고 논리적인 추론을 수행하는 과정을 실제에 가깝게 묘사할 수 있다는 장점을 가지고 있다. 그러나 아직도 프로그래밍 할 수 있는데는 한계가 있으며 더욱 실제에 근사하게 묘사하려는 노력은 계속되고 있다. 특히 많은 불확실한 요소가 내포되어 있고 전문가의 지식자체가 부정확한 경우가 많으며, 상황에 따라서 정보의 의미가 변화할 수 있는 동태적인 행태를 고려해야 하는 투자의사결정과 같은 복잡한 분야에서 프로그래밍할 수 있는 내용은 더욱 제한적일 수 밖에 없다.

이 논문도 퍼지퍼터덕션시스템에서 지식의 표현 및 추론과정을 좀더 실제에 가깝게 묘사하려는 노력의 일환으로 생각할 수 있다. 전문가의 지식을 IF-THEN 규칙으로 구성된 지식베이스로 구축하고 수시로 변환할 수 있는 규칙에 포함된 속성들의 중요도를 고려하여 평가함으로써 좀더 현실성있게 다중조건들을 평가하고 결합하여 결론을 도출할 수 있는 하나의 방법을 제시하였다. 지식베이스에 입력하여 추론결과를 얻으려는 입력자료(수집된 정보)와 지식베이스를 구성하는 규칙의 조건명제의 합치정도(DM)가  $[0, 1]$  범위값을 가지므로 하나의 규칙내에 있는 조건명제들의 중요도를 AHP(Analytic Hierachy Process)방법에 의한 가중치를 적용하여 정태적인 지식베이스에 동태적인 현실을 반영할 수 있도록 하였다.

이러한 방법은 지식분야에 관계없이 적용될 수 있으나 실용화를 위해서는 적용영역에 적합한 언어변수의 표현방법을 도출해야 한다. 즉 속성별 베이스값을 설정하고 언어변수값에 따른 멤버쉽값을 부여하는 멤버쉽함수를 도출해야하는데 이는 적용분야의 전문가들과 많은 시간의 지식도출(Knowledge Engineering)과정을 거쳐야 한다.

## 참 고 문 헌

1. Baldwin, J.F. and S.Q. Zhou, A Fuzzy relational inference language, *Fuzzy Sets and Systems*, 1982, 14.
2. Buchanan, B., Remarks delivered at North American Fuzzy Information Processing Conference II, Schenectady, NY, 1983.
3. Chandrasekaran, B., From Numbers to Symbols to Knowledge Structures : Pattern Recognition and Artificial Intelligence Perspective on the Classification Task. in : E.S. Gelsema and L.N. Kanal Eds., *Pattern Recognition in Practice-II*, North Holland Publishing Co., 1985.
4. Cho, S., O.K. Ersoy, and M. Lehto, An Algorithm to compute the degree of match in fuzzy systems, *Fuzzy Sets and Systems*, 49, pp 285-299, 1992.
5. Freksa, C., Linguistic description of human judgments in expert systems and soft sciences, in : M. Gupta and E. Sanchez Eds. *Approximate reasoning in decision analysis*, North-Holland, Amsterdam, 1982.
6. Garcia, J., J. Pazos, J. Rios and J. Yague, Metodology of linguistics evaluation in risk situations using fuzzy techniques, *Fuzzy Sets and Systems*, 48, pp. 185-194, 1992.
7. Ibrahim, A., and B.M. Ayyub, Multi-Criteria ranking of components according to their priority for inspection, *Fuzzy Sets and Systems*, 48, pp. 1-14, 1992.
8. Keen, P.G.W. and M.S. Scott Morton, *Decision Support System*, Addison-Wesley Publishing Co., Reading, Massachusetts, 1978.
9. Lesmo, L., L. Saitta and P. Torasso, Evidence com-

- bination in expert systems, *Int'l J. of Man-Machine Studies*, 22, pp. 307-326, 1985.
10. Mizumoto, M., Fuzzy inference using max composition in the compositional rule of inference, in : M.Gupta and E. Sanchez Eds., *Approximate reasoning in decision analysis*, North-Holland, Amsterdam, 1982.
  11. Newell, A. and H. Simon, *Human Problem Solving*, Prentice Hall, Englewood Cliffs, NJ, 1972.
  12. Prade, H., Approximate and plausible reasoning : the state of art., Communication at the IFAC Symposium on Fuzzy Information, Knowledge Representation, and Decision Analysis, July 1983, Marseille, France.
  13. Saaty, T., *The Analytic Hierachy Process*, McGraw Hill, New York, 1980.
  14. Saaty, T., Rank generation preservation, and reversal in the analytic hierachy decision process, *Decision Sciences*, Vol.18, No.2, pp. 157-177, 1987.
  15. Shafer, G. *A Mathematicl Theory of Evidence*, Princeton Univ. Press, 1976.
  16. Wenstop, Fred, Deductive verval models of organizations, *Int'l J. of Man-Machine Studies*, 8, 1976.
  17. Wenstop, Fred, Quantitative analysis with linguistic values, *Fuzzy Sets and Systems*, 4, 1980.
  18. Whalen, T. and B. Schott, Issues in fuzzy production systems, *Int'l J. Man-Machine Studies*, 19, pp. 57-71, 1983.
  19. Whalen, T. and B. Schott, Advances in linguistic processing, Report of the Decision Science Laboratory, Dept. of Decision Sciences, Georgia State Univ. Atlanta, GA, 1985.
  20. Zadeh, L.A., Fuzzy Sets, *Information and Control*, vol.8, pp. 338-353, 1965.
  21. Zadeh, L.A., The concept of a linguistic variable and its application to approximate reasoning I, II, III, *Information Sciences*, 8 : pp 199-249, pp 302-357, 9 : pp 43-80, 1975.
  22. Zadeh, L.A., Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems*, 1, pp 3-28, 1978.
  23. Zadeh, L.A., Test score semantics for natural languages and meaning representation via PRUF, in : B Rieger Eds., *Empirical semantics*, Brockmeyer, Bochum, West Germany, 1981.