

퍼지이론을 이용한 학점 전문가시스템연구

김 하 균¹⁾

배 경 울²⁾

I. 서 론

대학에서 교수님이 학생들의 성적을 산출 할 때 성적 등급과 학생 태도의 학점 반영을 퍼지전문가시스템을 통해 접근하며 문제의 접근 방향이 퍼지이론을 이용할 때 다른 비슷한 연구에 미칠 영향을 고려해 본다. 성적을 산출하는 방법은 교수님의 특성에 따라 다양하게 차이가 있으나 여기서의 고려 대상은 성적을 등급낼 때의 모순점과 개인적인 의견을 시스템에 어떻게 적용하는 것인가 하는 것이다. 이러한 문제점을 퍼지이론과 전문가시스템을 이용하여 해결하려고 한다. 전문가 시스템에서 필요 불가결한 것은 애매성과 모호성이다. 이러한 애매성의 문제를 퍼지이론을 이용하여 학점전문가 시스템에 응용할 것이다.

첫째로 성적을 줄 때에 가장 보편적으로 사용하는 것이 등급에 따라 성적을 산출하는 것이다. 즉 절대평가나 상대 평가는 90점이상 혹은 90%이상은 A, 80점에서 89점 혹은 80% 에서 89%은 B, 이와 같은 방법으로 나뉘어져서 중간 성적과 학기말 성적 등을 합하여 성적을 내는 것이다. 이 방법은 불합리한 점을 가지고 있다. 즉, 80점과 89점의 경우 9점의 차이에도 같은 성적 등급인 B이지만 89점과 90점의 경우 1점의 차이에도 B와 A의 성적 등급의 차이가 난다. 이러한 모순점을 해결하기 위하여 퍼지이론을 이용하고자 한다. 즉 90점의 경우 B와 A될 가능성³⁾은 양쪽 모두 0.5인 것이다. 즉, 89점의 경우는 기존의 점수인 경우는 B가 되겠지만, 퍼지이론을 사용할 경우는 A가 될 가능성이 0.4이고 B가 될 가능성이 0.6이 된다(가능성 0.4와 0.6은 퍼지함수의 폭에 따라 차이가 있으며 퍼지함수의 폭은 의사결정자에 의해 정해진다). 이와 같이 모든 점수는 일방적인 성적을 갖는 것이 아니라 양쪽의 모든 가능성을 갖고 있다는 것이다. 다른 한편으로는 91점의 경우는 A가 될 가능성이 0.6이 되고 B가 될 가능성도 0.4가 될 것이다. 이것은 지금까지의 등급에서 벗어나 양쪽의 가능성을 고려해 보는 것이다.

두번째로 이 연구의 고려점은 성적을 산출할 때 교수의 의견을 성적에대한 반영도이다. 즉 성적은 점수만 가지고 이루어지지 않는다. 즉, 중간고사와 학기말에서 성적이 A인 경우에도 출석이 70%가 넘지 않으면 학교 규정상 F가 되는 것이다. 이러한 경우는 간단하

1) 부경대학교 경영정보학과

2) 상명대학교 정보처리학과

3) 퍼지이론에서는 확률(probability)대신 가능성(Possibility)를 사용함.

한국과학기술원, 퍼지이론 및 응용, 홍릉과학출판사, pp 8-1~8-32

지만 실제로 출석이 75%에다 과제물도 제출하지 않은 학생이 시험 성적이 좋다는 이유만으로 A를 받는 것은 불공평하다. 즉 이러한 문제를 해결하기 위하여 퍼지전문가 시스템의 도입이 필요한 것이다. 교수님의 의견 자체가 계량화되지 않았으므로 수량화 할 수 있는 퍼지이론과 전문가 시스템에서는 학생에 대한 교수의 의견을 규칙(Rule base)화 함으로써 성적에 반영하는 것이다.

위의 두 가지 문제점에 초점을 맞추어 시스템화함으로써 성적을 좀더 정확하고 성적 산출에 대한 교수님을 시간을 절약하는데 그 목적이 있으며 지금까지 많은 시스템이 입력 등급(input Grade)에 의해 출력(Output)이 결정되는데, 입력 등급의 양쪽을 모두 고려함으로써 그러한 모순점을 줄일 수가 있다. 이 연구의 복잡성을 줄이기 위해 시험의 성적은 중간 성적과 학기말 성적 두번으로 정했다. 성적은 A, B, C, D, F의 다섯 등급으로 나누었으며, 교수님의 학생에 대한 의견은 좋다(Good), 보통이다(OK), 나쁘다(Bad), 매우나쁘다(Very Bad)의 네등급으로 나누었다.

II. 퍼지전문가 시스템

현재 전문가 시스템의 구축 방향은 애매한 정보의 처리, 추론능력에 있다고 할 수 있다. 이러한 애매한 정보를 퍼지화 함으로써 정확한 전문가 시스템의 구축이 필수적이다. 따라서 퍼지전문가시스템의 구축은 급격히 증가하고 있으며 성공사례도 다양하다. 전문가시스템의 모호한 지식으로 표현되는 것으로는 비결정론(non-determinism), 복합의미(multiple meanings), 불완전성(incompleteness), 불확실성(uncertainty), 퍼지성(fuzziness) 등이다. 이중 불확실성과 퍼지성은 전문가시스템에서 퍼지이론을 응용하여 좋은 결과를 얻을 수 있는 분야이다. 성공한 예로는 버스 노선의 전문가가 고객 수요, 교통 사정 등을 고려한 버스 다이어그램 편성 전문가시스템, 달 착륙에 대한 시나리오 작성후 거리와 속도에 대하여 퍼지화한 달착륙전문가 시스템, 건물 안에 사람의 인원을 자동적으로 탐지하는 감지장치를 이용한 인원 탐지 전문가 시스템, 건물의 손상평가를 위한 SPERIL 시스템, 퍼지추론을 이용한 의료진단전문가 시스템은Sanche(1982)를 비롯한 Adlassing(1986), Tazaki(1988) 등 연구가 상당히 진행 중이다. 이러한 예는 전문가 시스템에 애매한 정보를 퍼지화 함으로써 성공한 퍼지전문가시스템들이다.

퍼지이론에 대하여서는 많은 연구가 진행되어 왔으나 기본적으로 퍼지성(Fuzziness)이라는 것은 단어의 의미나 개념을 정의할 때 생길 수 있는 모호함이다. 즉 젊은 사람, 높은 온도, 작은 수, 빠른 속도 등 표현에서 포함되는 불확실성인 것이다. 이러한 퍼지성으로부터 인간의 느낌, 상식, 언어 표현, 심리 등을 모델화 할 수 있는 것이 퍼지논리이다. 간단한 예로써 퍼지자동차 브레이크시스템을 보면 앞차와의 거리 (가깝다, 보통이다, 멀다)와 속도(빠르다, 보통이다, 느리다)를 입력 변수로 9개의 규칙(rule)을 통해 퍼지추론함으로써 브레이크의 정도(세기, 보통, 약하게)을 자동적으로 출력할 수 있는 것이다.

전문가시스템은 다음과 같은 특징이 있다. 첫째로 전문가적인 의견으로써 문제해결에 도움을 주어야 한다. 즉 최상의 지식을 갖춘 상급전문가가 있어야 한다. 둘째로 전문가 시

시스템은 모델링 능력이 있어야 한다. 즉 문제해결과 그 상황에 맞는 해답을 주기위한 시스템화가 되어야 한다. 문제의 결과에서는 문제해결을 위한 설명도 있어야 한다. 위와 같은 전문가 시스템은 단순한 프로그램이 아닌 전문가의 문제를 효율적으로 해결하기 위한 다양한 이론과 기술이 제공되어야 하고 모델화 되어야 한다. 전문가 시스템이 가장 효과적인 분야는 접근 방향이 어렵거나 수식화가 불가능한 분야이다. 이러한 경우 문제의 지식이 분명치 않다는 것이다. 다시말 하면 문제의 지식이 모호하다는 것이다. 이러한 모호함 또는 모호한 지식을 컴퓨터가 이해할 수 있도록 적절한 형태를 취해야만 한다. 이러한 모호함과 불확실성에 대한 여러가지 이론(퍼지이론, Dempster-Shafer이론, Bayesian이론)이 있으며 전문가시스템에서 다양하게 적용되어 왔었다. 여기서는 이러한 모호함 즉, 학생에대한 교수의견의 모호함과 성적 등급의 불합리성을 퍼지화함으로서 해결하고자 한다

Ⅲ. 퍼지학점전문가 시스템의 모델설정

퍼지 학점전문가 시스템을 구축하기 위하여는 다음과 같은 단계의 시스템 설계를 따른 것이다.

1) 입력데이터의 범위와 명칭을 확정한다;

입력 변수는 복잡성을 줄이기 위하여 중간 성적, 학기말 성적, 교수의 의견 등 세 가지 변수를 사용했다. 성적의 등급은 A, B, C, D, F의 다섯 등급으로 교수의 의견은 좋다, 보통이다, 나쁘다, 매우 나쁘다 등 네 등급으로 나뉘었다. 범위와 명칭은 <표 1>과 <표 2>와 같다.

<표 1> 중간고사 와 학기말 성적 등급

학 점	F	D	C	B	A
점 수	0 — 59	60 — 69	70 — 79	80 — 89	90+

<표 2> 교수의견 등급

교수의견	매우 나쁘다	나쁘다	보통이다	좋다
학생태도4)	0 — 69	70 — 79	80 — 89	90+

2) 출력데이터의 범위와 명칭을 확정한다;

출력데이터의 경우 최종 성적을 나타내며 성적의 결과를 A, B, C, D, F 의 다섯 등급으로 나눈다. 범위와 명칭은 <표 3>과 같다.

4) 출석, 수업태도, 과제물 제출 등을 고려하였다.

<표 3> 최종 성적

학 점	F	D	C	B	A
점 수	0 — 59	60 — 69	70 — 79	80 — 89	90+

3) 입력 및 출력에 대한 퍼지소속함수(Fuzzy Membership Function)을 작성한다.

<표 1> <표 2> <표 3>에 대한 퍼지소속함수를 작성한다. <그림 1>은 성적, <그림 2>는 교수의견에 대한 퍼지입력함수이며, <그림 3>은 최종 성적에 대한 퍼지출력함수이다.

4) 입력과 출력데이터를 AS Something I THEN Something II으로 지식 베이스를 만든다.

<표 4> <표 5>와 같이 규칙 베이스를 만든다. <표 4>는 중간 성적과 학기말 성적 25개의 규칙으로 되어 있으며 <표 5>는 최종 성적과 교수의견의 20개 규칙으로 되어 있다. 두개의 규칙 베이스를 합하여 최종적으로 퍼지학점전문가 시스템은 45개의 규칙을 갖는다.

<표 4> 중간 성적과 학기말 성적 규칙5)

중간성적 학기말	F	D	C	B	A
A	C	B	B	A	A
B	C	C	B	B	B
C	D	C	C	C	B
D	D	D	D	C	C
F	F	F	D	D	C

<표 5> 최종 성적과 교수의견 규칙6)

최종성적 교수의견	F	D	C	B	A
좋다	D	C	B	A	A
보통이다	F	D	C	B	A
나쁘다	F	F	D	C	B
매우 나쁘다	F	F	F	F	F

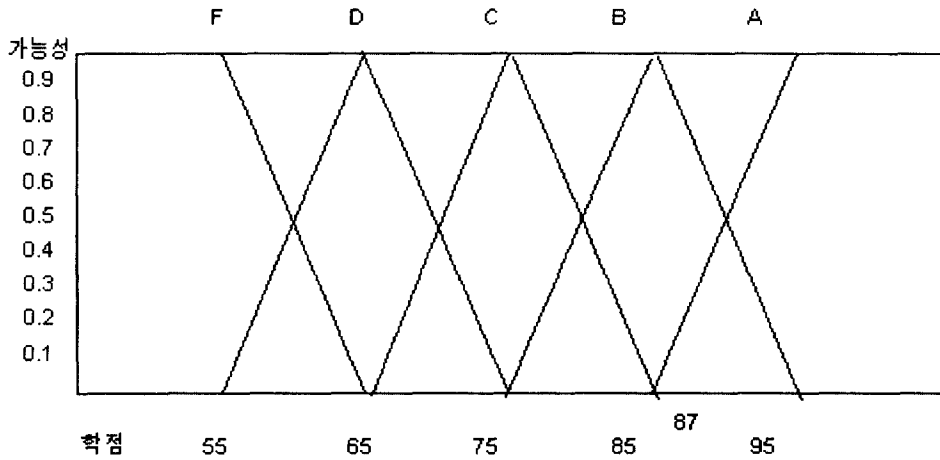
5) 25개의 규칙 행렬(Rule Matrix)로 이루어져 있음

예) 중간 성적 “F”이고 학기말 성적 “A”이므로 성적은 “C”이다.

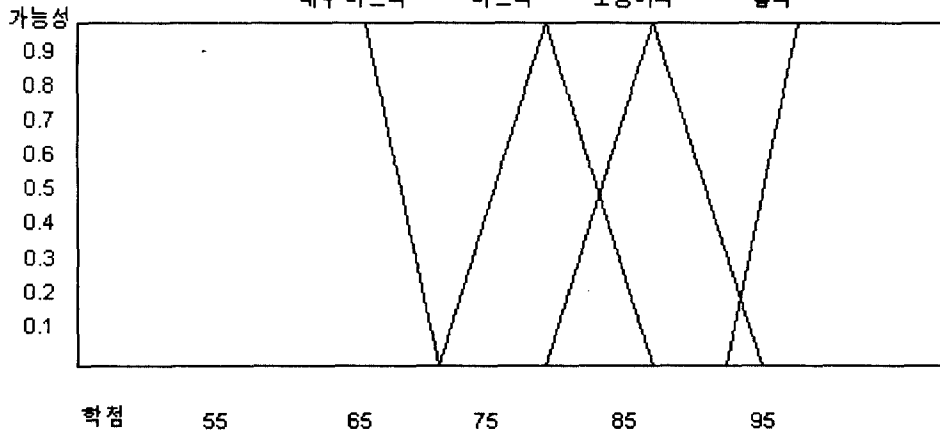
6) 20개의 규칙 행렬(Rule Matrix)로 이루어져 있음

예) 최종 성적 “F”이고 교수의견 좋다이므로 성적은 “D”이다.

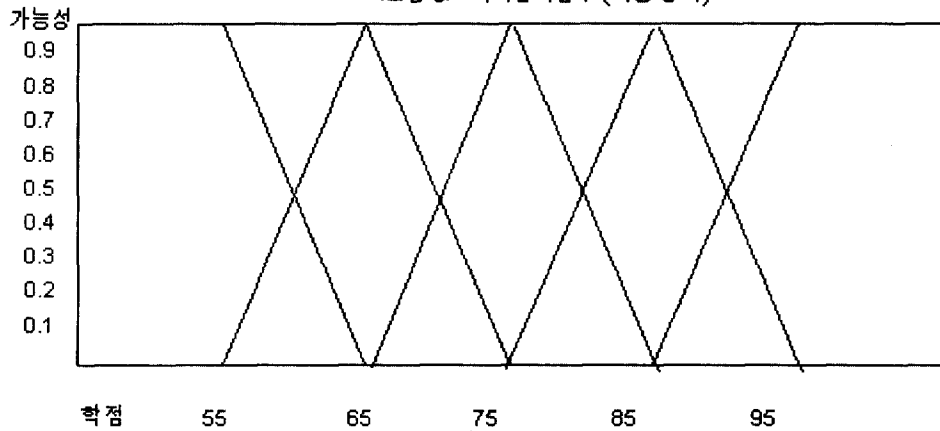
<그림 1> 퍼지 입력함수 (중간성적, 학기말 성적)



<그림 2> 퍼지 입력함수 (교수 의견)
매우 나쁘다 나쁘다 보통이다 좋다



<그림 3> 퍼지 출력함수 (최종 성적)



5) 탈퍼지화(Defuzzified)함으로서 입력데이터와 지식베이스를 기준으로 작동할 추론엔진을 결정한다.

퍼지함수의 논리합에 대한 결과를 탈퍼지화 하기 위하여 퍼지논리 합의 값에서 중앙값(Center Value)으로 정하였다. <그림 5-1> <그림 5-2>은 논리합에 대한 예제이다.

<그림 5-1>의 경우는 입력함수로서 학기말 성적 (91점), 중간 성적 (81점)을 받은 경우이다. 학기말 성적 91점은 A함수에서 “A”될 가능성은 0.6, B함수에서 “B”가 될 가능성은 0.4이다. 즉, 퍼지입력함수 값은 A0.6, B0.4이다. 중간 성적 81점은 B함수에서 “B”될 가능성은 0.6, C함수에서 “C”가 될 가능성은 0.4이다. 즉, 퍼지입력함수 값은 B0.6, C0.4이다. 이것은 학점전문가 시스템의 규칙베이스(Rule Base)에 의해 네가지 규칙을 만들 수 있다.

- ① 최종 성적 “A0.6”이고 중간 성적 “B0.6” 이므로 최종 성적은 “A0.6”이다.
- ② 최종 성적 “A0.6”이고 중간 성적 “C0.4” 이므로 최종 성적은 “B0.4”이다.
- ③ 최종 성적 “B0.4”이고 중간 성적 “B0.6” 이므로 최종 성적은 “B0.4”이다.
- ④ 최종 성적 “B0.4”이고 중간 성적 “C0.4” 이므로 최종 성적은 “B0.4”이다.

따라서 <그림 5-1>에서의 윗그림처럼 논리합은 A0.6, B0.4의 합이며 탈 퍼지화한 논리합의 중앙값은 “87”이다.

<그림 5-2>의 경우는 입력함수로서 최종 성적(84점), 교수의견 (93%)을 받은 경우이다. 최종 성적 84점은 B함수에서 “B”될 가능성은 0.9, C함수에서 “C”가 될 가능성은 0.1이다. 즉, 퍼지입력함수 값은 B0.9, C0.1이다. 교수의견 93%은 좋다함수에서 “좋다”가 될 가능성은 0.2, 보통함수에서 “보통이다”가 될 가능성은 0.2이다. 즉, 퍼지입력함수 값은 “좋다0.2”, “보통이다0.2” 이다. 이것은 학점전문가 시스템의 규칙베이스(Rule Base) 의해 네 가지 규칙을 만들 수 있다.

- ① 최종 성적 “B0.9”이고 교수의견 “좋다0.2” 이므로 성적은 “A0.2”이다.
- ② 최종 성적 “B0.9”이고 교수의견 “보통이다0.2” 이므로 성적은 “B0.2”이다.
- ③ 최종 성적 “C0.1”이고 교수의견 “좋다0.2” 이므로 성적은 “B0.1”이다.
- ④ 최종 성적 “C0.1”이고 교수의견 “보통이다0.1” 이므로 성적은 “C0.1”이다.

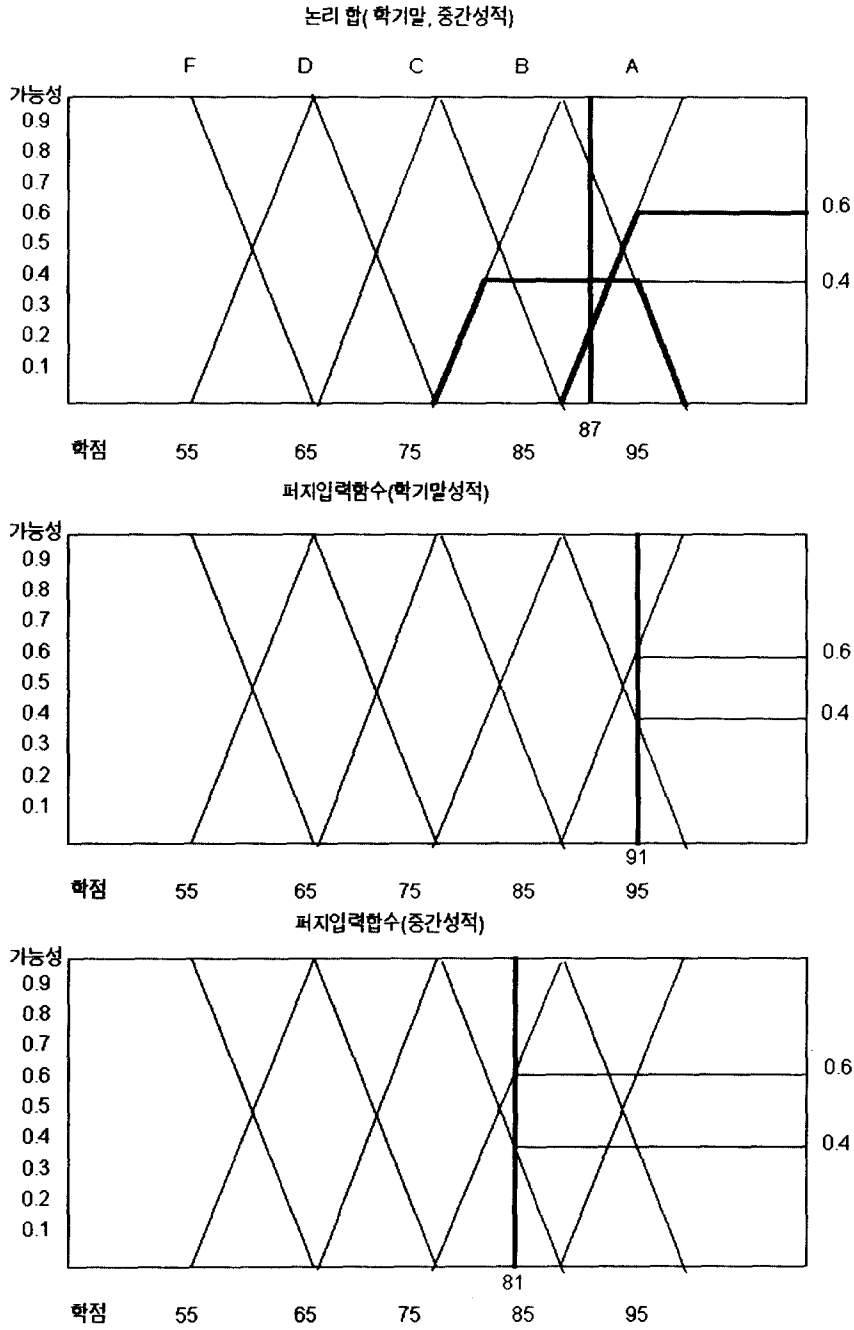
따라서 <그림 5-2>에서의 윗그림처럼 논리합은 A0.2 , B0.2 , C0.1의 합이며 탈 퍼지화한 논리합의 중앙값은 “85”이다.

6) 퍼지학점전문가 시스템의 정확도를 시험한다.

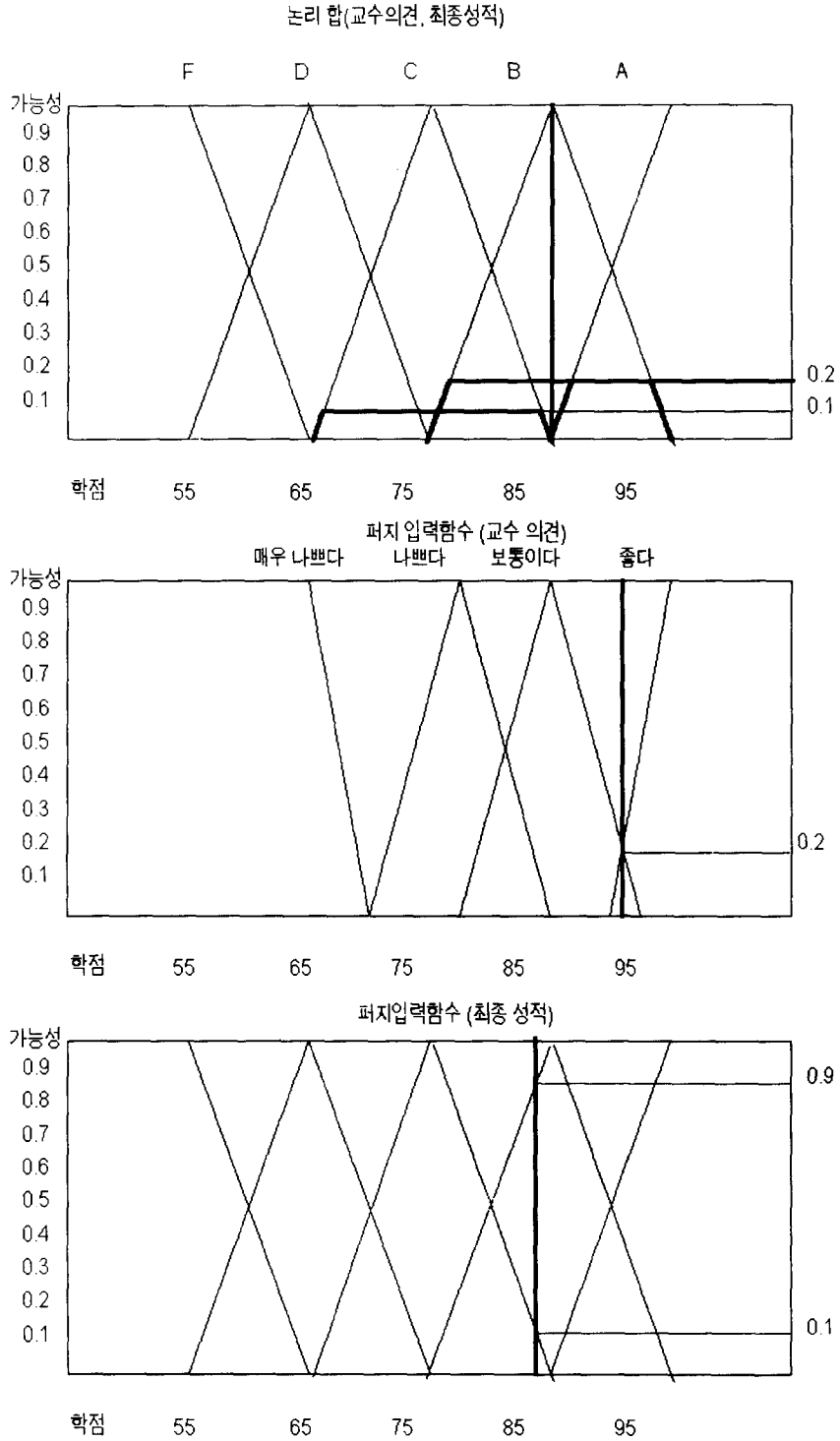
퍼지학점전문가 시스템의 정확도를 시험하기 위하여 세 명의 전문가(교수 A, B, C)로 부터 성적을 받은 후 퍼지학점 전문가 시스템에 입력하여 몇 퍼센트의 정확도를 갖는지를 측정한다. 결과는 <표 6>과 같다.

7) 세명의 교수로 구성되었으며 이 논문의 성격을 정확히 인지하고 있음

<그림 5-1> 학기말 성적과 중간고사의 탈퍼지화



<그림 5-2> 교수의견과 최종 성적의 탈퍼지화



<표 6> 퍼지학점전문가 시스템의 정확도

구 분	전문가 A	전문가 B	전문가 C
퍼지학점전문가시스템	95 %	90 %	88 %

IV. 결 론

퍼지학점전문가시스템은 교수가 학점을 산출함에 있어서 성적 등급의 문제점과 교수의 의견을 학점에 어떻게 반영하여 정확한 성적을 반영하는 전문가 시스템을 구축하는데 목적이 있다. 이러한 두 가지 문제점을 해결하기 위하여 전문가시스템에 퍼지이론을 도입했으며 최근에 전문가시스템에 퍼지이론을 이용한 시도는 다양하며 널리 사용되고 있다. 퍼지이론을 사용한 전문가시스템의 경우 전문가시스템 자체보다 성공사례가 많다.

퍼지학점전문가 시스템을 세명의 전문가에 의해 평가해 본 결과 최소 88%에서 최고 95%까지의 높은 적중률을 보이고 있다. 따라서 이 연구결과는 성공적이라고 할 수 있다. 전문가 시스템으로는 성공적이라고 할 수 있는 것은 성적이라는 특수성 때문에 의사결정의 큰 비중이 정량화 되어 중간고사와 학기말 점수에 의존하며 전문가(교수)의 의견 (학생의 수업 태도나 과제물 제출)이 많이 반영되지 않았기 때문이다.

이러한 점에서 볼 때 이 연구의 공헌은 전문가의 의견에 대한 학점의 반영보다도 학점을 등급 매김에 있어서 단편적인 면이 아닌, 90점의 경우 “A”가 될 가능성이 50%, “B”가 될 가능성이 50%인 양쪽의 가능성을 모두 고려 한데 그 의의가 있다고 할 수 있다.

문제점으로는 모든 퍼지이론이 그렇듯이 퍼지함수가 개인의 의사결정에 의해 결정되어 진다는 것이다. 이 논문의 경우 학점에는 퍼지함수화가 문제되지 않지만 교수의 의견은 개인의 의사결정이 중대한 영향을 결과에 반영됨으로 정밀한 자료조사 후 결정되어야 할 것이다. 이 연구를 더욱 발전시키기 위하여는 입력 변수(중간 성적, 학기말 성적, 교수의 견) 사이에 가중치를 줌으로써 모든 입력 변수를 동시에 고려하는 것이다. 이와 같은 방법은 퍼지함수와 규칙 매트릭스를 더욱 복잡하게 하고, 가중치 문제 등 개인의 의사결정이 더욱 커지는 것 또한 문제이다.

References

- 김도현, 핵심 퍼지시스템 이론 및 응용서, 에드텍.
박민용, 최항식, 퍼지시스템의 응용입문, 법문사.
엄정국, 퍼지이론(기초와 응용입문), 박영사.
- Adlassing, K.P. (1986), Fuzzy Set Theory in Medical Diagnosis, IEEE Trans. S.M.C., SMC-16, 2, pp 280-265.
- Cox, E.(1991), Fuzzy logic Clarified(Application of Fuzzy Set Theory to Information Systems), Computer World, 11, pp 69-71
- Hisdal, E.(1988), The Philosophical Issues Rasied by Fuzzy Set Theory, Fuzzy Sets and Systems, 25, pp 349-356.
- Negoita, C.(1987), Simulation, Knowledge-based Computing and Fuzzy Statistics, NewYork:Van Nostrand Reinhold.
- Prade, H.(1985), A Conceptual Approach to Approximate and Plausible Reasoning with Application to Expert Systems, IEEE Transactions, PAMI, 3, pp 774-779.
- Sanchez, E. (1982), Linguistic approach in fuzzy logic of W.H.O. classification of dyslipoproteinemias, Fuzzy Set and Possibility Theory-Recent Development, Yager ed. Pergamon, pp 582-588.
- Tazaki, E. (1986), Development of Automated Health Testing and Services System via Fuzzy Reasoning, Proc. IEEE Int. Conf. on SMC, pp 342-346.
- Whalen, T (1983), Issues in Fuzzy Production Systems, International Journal of Man-Machine Studies, 19, pp 57-81.
- Zadeh, L.(1968), Probability Measure of Fuzzy Events, Jounal of Mathematical Analysis and Application, 23, pp 421-427.