

다양한 퍼지 환경을 갖는 지능형 교수 시스템의 학습 성취도 평가 모듈 설계*

원 성 현**

〈목 차〉

I. 서론	IV. 모의 실험과 결과의 토론
II. ICAI의 기본 구조와 평가	1. 모의 실험 대상
1. ICAI의 기본 구조	2. 학습자의 응답 결과
2. 평가와 지능적 평가	3. 지능적 평가
3. 지능적 평가를 위한 기준 연구	4. 결과의 토론
III. 학습 성취도 평가 모듈	V. 결론
1. 응답 정확도 산출	참고문헌
2. 평가 함수와 지능적 평가	Abstract

I. 서 론

CPU 가격의 하락과 컴퓨터 제작 기술의 눈부신 발달은 개인용 컴퓨터의 보급을 촉진시키는 결정적인 계기가 되었고, 여기에 최근 멀티미디어라는 새로운 컴퓨팅 기술이 도입되면서 이전까지 학습지, 카세트 테이프, 비디오 테이프 등 단일 미디어를 기반으로 하는 학습 보조 시스템이 멀티미디어 기반의 개인용 컴퓨터를 이용한 보조 교육 시스템(Computer Assisted Instruction : 이하 CAI로 약칭)으로 과감하게 이동하고 있다(박성익, 1988). CAI에 대한 연구는 전통적으로 교육학 분야에서 수행했으나 최근 컴퓨터 공학 분야에서도 활발히 연구하고 있는 추세이다.

CAI는 1970년대 초반부터 연구되어 왔고, 최근에는 인공지능(artificial intelligence) 기

*이 논문은 1996년 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

**지산대학 전산정보처리과 조교수

법을 도입한 지능형 CAI(Intelligent CAI: 이하 ICAI로 약칭)도 다수 개발되어 학습자들을 위한 보조 학습 도구로 많이 활용되고 있다(박성익, 1988).

CAI의 구성 요소로는 영역 전문가 모듈(domain expertise module), 학습자 모델 모듈(student model module), 학습 지도 모듈(tutoring module), 상호 접속 모듈(interface module) 등이 있다. 그리고 이들 모듈은 각각 세부 기능들로 구조화되어 있는데, 그 중 영역 전문가 모듈에서의 평가(evaluation) 기능은 다른 기능들과 비교했을 때 특히 그 중요성이 강조되고 있다. 왜냐하면 학습자의 응답에 대한 적절한 평가가 이루어져야만 학습자에게 최적의 학습 전략을 수립하여 제시할 수 있고, 이를 통해 최고의 학습 효율을 유도할 수 있기 때문이다(Kearsley, 1987).

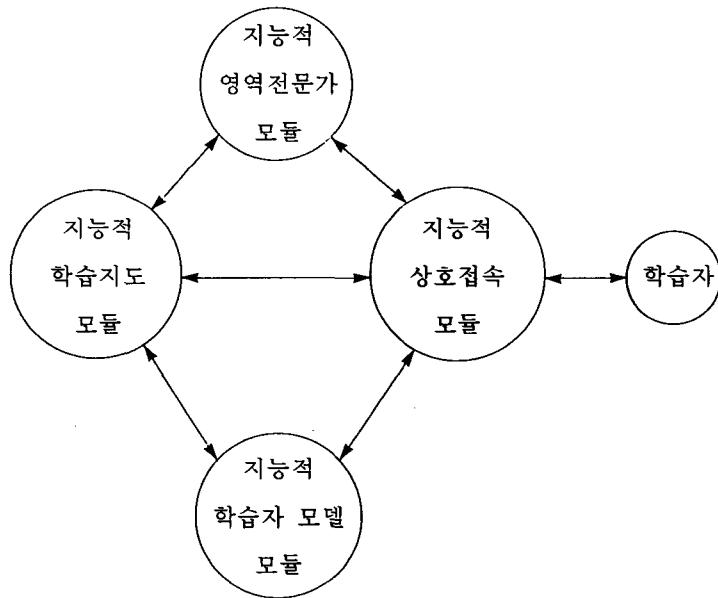
전통적으로 사용되어 왔던 등급에 의한 학습 성취도 평가(learning achievement evaluation)는 인간 교사(human teacher)의 잠재적 평가 방식과 상이하다는 단점이 있음에도 불구하고 절차가 용이하다는 장점 때문에 최근까지도 보편적으로 사용해 왔던 방식이다. 그러나 초등학교를 중심으로 학생들에 대한 평가를 전통적인 수·우·미·양·가식 평가를 지양하고 언어적 표현 방식에 의한 평가를 지향하게 됨으로써 평가 요소에 대한 반영 방식이 새로운 이슈가 되고 있다.

본 연구에서는 전통적인 평가 방식이 갖고 있는 문제점을 해결하고자 한다. 즉, 전통적인 평가 방식은 평가 문항에 내포되어 있는 문항의 복잡성, 중요성, 난이성 등을 크게 고려하지 않았다. 그러나, 이들 요소들을 적절하게 고려해야만 정확한 평가가 이루어질 수 있다는 점이 여러 전문가들에 의해 제기되어 있으나 이들에 대한 연구는 많이 이루어지지 않고 있는 실정이다. 따라서, 모호한 환경 하에서의 효과적인 평가 측도를 제공하고 있는 퍼지 이론(Yager, 1981; Zadeh, 1976)을 이용하여 이들과 같은 다양한 퍼지 환경 하에서의 적절한 평가를 이루고자 한다. 이를 위해, 퍼지 이론의 퍼지화 함수를 이용한 평가 식을 만들고 이 식에 의해 다양한 퍼지 환경을 수용하는 자동화된 학습 성취도 평가 모듈을 설계하여 학습자에게 언어적인 평가 결과를 제공할 수 있도록 한다.

II. ICAI의 기본 구조와 평가

1. ICAI의 기본 구조

I 장에서 언급했듯이 CAI의 구성 요소로는 영역 전문가 모듈(domain expertise module), 학습자 모델(student model module), 학습 지도 모듈(tutoring module), 상호 접속 모듈(interface module) 등이 있다. ICAI는 이들 각 모듈에 지능적 기능을 추가한 것으로 생각할 수 있다. <그림 1>은 ICAI의 기본 구조를 나타낸다.



<그림 1> ICAI의 기본 구조

지능적 영역 전문가 모듈(intelligent domain expertise module)은 학습자에게 가르치고자 하는 영역에 대한 지식으로 구성되어 문제를 생성하고 학습자의 응답에 대한 옳고 그름을 평가한다. 또한, 기존의 지식으로부터 새로운 지식을 유도하여 학습시키는 기능도 포함된다. 지능적 학습자 모델 모듈(intelligent student model module)은 학습자에 관한 지식 및 행동 등 학습 과정에서 학습자가 보인 학습 성향 등으로 구성되고 학습자가 학습 과정에서 범할 수 있는 오류, 학습자의 학습 상의 문제점들을 진단할 수 있는 기능

을 갖고 있다. 지능적 학습 지도 모듈(intelligent tutoring module)은 다양한 종류의 학습 전략과 학습 상의 문제점들을 진단하고 처방하는 데 필요한 지식들로 이루어져 있고 지능적 상호 접속 모듈(intelligent interface module)은 학습자와 앞의 세 구성 요소들이 효과적으로 정보를 교환할 수 있도록 하는 기능으로 구성된다(박성익, 1988).

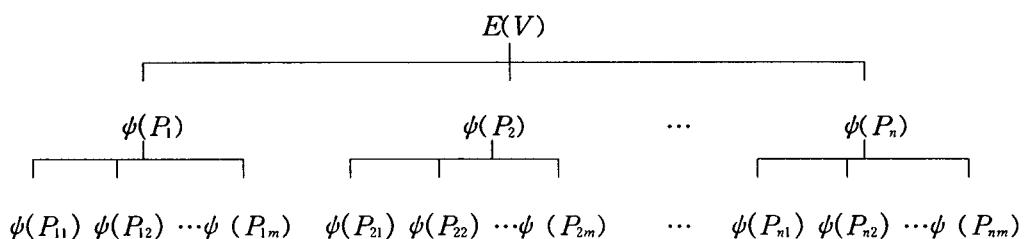
2. 평가와 지능적 평가

평가란 통상 교육 평가 혹은 학습 성취도 평가의 줄임말로서 사전적으로 사물 또는 그 속성에 대한 가치 판단이라고 정의할 수 있다. 다시 말하면, 학습된 내용에 대해 성취된 정도를 정해진 척도에 의해 판단하는 것을 말한다(변창진, 1986).

일반적으로 학습 성취도의 평가는 개개의 평가 요소에 대해서 각각의 요소가 더욱 하위의 평가 요소를 갖는 계층 구조로서의 종합적인 평가가 될 수 있다. 따라서, 학습 상태의 평가를 E 라 하면 E 는 하위의 평가 요소 $P_i (1 \leq i \leq n)$ 의 평가 함수 $E = \psi(P)$ 로 결정된다(山下 元・清水 誠一, 1993).

또한, E 의 평가 요소 $P_i (1 \leq i \leq n)$ 은 그 하위의 평가 요소 $P_{ij} (1 \leq j \leq m)$ 의 평가 함수 $P_i = \psi(P_{ij})$ 로 결정된다.

따라서, E 의 평가 계열은 〈그림 2〉와 같은 구조로 생성된다.



〈그림 2〉 평가계열의 구조도

지능적 평가란 이와 같은 평가 방식에 인간 교사의 언어적 평가 방식을 혼합한 것이다. 즉, 교사가 학생을 평가할 때, 100점 만점에 몇점을 얻었으니 ‘잘했다’, ‘보통이다’, ‘못했다’ 등으로 평가하곤 한다. 또한, 학생이 해결해야 할 문제와 관계해서는 ‘복잡하다’, ‘보통이다’, ‘간단하다’ 등의 복잡도에 관한 요소가 있을 수 있고, ‘중요하다’, ‘보통이다’, ‘중요하지 않다’ 등의 중요도에 관한 요소가 있을 수 있고, ‘어렵다’, ‘보통이다’, ‘쉽다’ 등의 난이도에 관한 요소가 있을 수 있다. 이 때 복잡도, 중요도, 난이도 등을 언어 변수 (linguistic variable)라고 하고, ‘중요하다’, ‘보통이다’, ‘쉽다’ 등을 폐지 변수(fuzzy variable)라고 한다. 이와 같이 어떤 정형화된 식이나 규칙에 의해 평가하지 않고 출제된 문제에 여러 언어 변수들이 포함되어 있고, 이를 언어 변수에 폐지 변수가 할당되어 있으며 이들을 통해 언어적 평가 결과를 산출하도록 하는 평가 방식을 지능적 평가라고 한다.

3. 지능적 평가를 위한 기존 연구

평가에 대한 연구는 CAI뿐 아니라 전통적인 교육학에서도 많이 연구되어 왔다. 그러나, ICAI에서의 지능적 평가에 대한 연구는 최근에 와서야 비로소 그 연구의 필요성이 대두되게 되었다. 국내에서는 서대우 등에 의해 지능형 저작 시스템에서 언어 변수를 이용한 학습자 모듈이 제안(서대우 외 공동연구, 1993)되기도 했고, 이민숙 등(1992)에 의해 학습자 수준과 문제 난이도에 따라 시간을 부여하는 학습자 모듈도 제안되었다. 앞서 언급한 중요성이나 복잡성 등을 고려한 폐지 평가에 대한 연구들이 원성현 등(1994, 1996)에 의해 진행되었으나 이들은 각각의 폐지 환경을 통합하지 못했다는 문제점이 있다. 또한, Law(1996)에 의해 폐지 수(fuzzy number)를 이용한 교육 등급 시스템이 제안되었으나 이 제안은 처리 과정이 지나치게 복잡한 수학적 처리 절차를 요구하고 있고 더욱이 다양한 폐지 환경을 통합하지 못했다는 문제점이 있다. 일본을 중심으로 폐지 평가에 대한 연구도 진행되었는데 그 시작은 郁雄(1991)에 의해서 였으며, 慶三 등(1991)은 학습자의 응답 시간에 대한 행렬을 해석하는 방법으로 폐지 평가 방법을 제안하고 있다. 元 등(1993)은 서예와 같은 주관적 평가 요소가 많이 포함되어 있는 경우에 대한 폐지 평가 방법을 제안했으며, 龍彌 등(1994)은 신경망을 이용한 방법도 실험적으로 제안하고 있다. 그러나, 이를 방법 모두 실험적인 제안에 그쳤거나 다양한 폐지 환경의 통합에 대해 해결하지 못하고 있다. 따라서, 최근의 교육계에서의 평가 요구를 반영할 수

있는 평가 방법으로 다양한 폐기 환경을 통합하여 처리할 수 있는 평가 모듈의 개발이 절실히 요구되고 있다.

III. 학습 성취도 평가 모듈

1. 응답 정확도 산출

1.1 제한 시간만 있는 경우의 응답 정확도

임의의 문제 영역(P)에 관계된 몇 개의 문항(P_i)이 식 (3)과 같이 주어졌고, 각 문항에는 몇 개의 부문항(P_{ij})이 다음 식 (4)와 같이 주어졌다면 이들에 대한 응답 정확도 $COR(P)$ 는 다음 식 (5)로 정의한다.

단, P 는 문제 영역이고, P_i 는 문제 영역 내의 임의의 문항이다.

$$= \bigcup_{i=1}^n \{P_i, \sum_{j=1}^m (\mu P_{ij} \times \mu T_{ij})\} \quad \dots \dots \dots \quad (5)$$

단, μP_i 는 i 번째 문항의 j 번째 부문항의 정답 여부에 대한 귀속도로 정답이면 1, 아니면 0이고, Σ 는 대수적 합, \cup 은 집합, \times 는 대수적 곱, μT_j 는 P_j 를 해결하는 데 소요된 시간에 대한 귀속도를 의미한다.

또한, μT_i 는 다음 식 (6)과 같은 변형된 시그모이드 함수로 구한다.

단, v 는 P_i 를 해결하는 데 소요된 시간, α 는 문제 해결에 허용된 하한 시간, γ 는 문제 해결에 허용된 상한 시간, β 는 $(\alpha+\gamma)/2$ 를 의미한다.

1.2 중요도를 고려한 경우의 응답 정확도

중요도란 그 문항이 학습자가 문제 영역을 이해하는 데 얼마나 중요한 비중을 차지하는가를 나타내는 척도이다. 그래서 중요성이 높은 문항 즉, 중요도가 ‘중요하다’인 문항을 해결한 경우에는 가중치를 주고, 상대적으로 중요성이 낮은 문항 즉, ‘중요하지 않다’인 문항을 해결한 경우에는 가중치를 빼는 방법을 생각할 수 있으며, 중요성이 보통인 문항 즉, 중요도가 ‘보통이다’인 문항에 대해서는 본래의 정확도를 그대로 두는 것이 타당하다. 가중치를 주기 위해선, 폐지 이론의 팽창(dilation)기법을 이용하고, 가중치를 빼기 위해선, 역시 폐지 이론의 압축(concentration)기법을 이용하기로 한다. 이와 같은 방법으로 중요도가 고려된 경우의 응답 정확도 산출식 ICOR을 식 (7)에 나타냈다.

단, k 는 가중치 인자로 ‘중요하다’인 경우는 0.5, ‘보통이다’인 경우는 1, ‘중요하지 않다’인 경우는 2로 정의한다.

1.3 복잡도를 고려한 경우의 응답 정확도

복잡도란 ‘어렵거나 쉽다’는 개념과는 다르다. 문항이 복잡하다는 것은 학습자의 실력과 전혀 무관하지는 않으나 통상 제한 시간이 충분히 주어지면 성공적으로 응답할 수 있는 경우가 많다. 따라서, 복잡도와 관계된 문항은 문제 해결 허용 시간을 조정함으로써 해결할 수 있다. 복잡도가 ‘복잡하다’, ‘보통이다’, ‘단순하다’의 세 종류로 구분되었을 때, 각 문항에 대해 응답한 모든 학습자의 응답 시간에 대한 표준편차(σ)를 구하여 복잡도가 ‘복잡하다’인 경우는 제한 상한 시간을 표준편차만큼 늘려주고, ‘단순하다’인 경우는 제한 상한 시간을 표준편차만큼 줄인다. 그리고, ‘보통이다’인 경우는 최초의 제한 시간을 그대로 유지하는 방법을 생각한다. 이와 같은 방법으로 복잡도가 고려된 경우의 응답 정확도 산출식 CCOR을 식 (8)에 나타냈다.

$$\text{CCOR}(P) = \{(P_1, \text{COR}(P_1)), (P_2, \text{COR}(P_2)), \dots, (P_n, \text{COR}(P_n))\}$$

이 때, $\mu T'$ 는 다음 식 (9)로 구한다.

$$\mu T'_{ij} = \begin{cases} 1 & : v \leq \alpha \\ 1 - 2((v - \alpha)/(\gamma' - \alpha))^2 & : \alpha < v \leq \beta' \\ 2((v - \gamma')/(\gamma' - \alpha))^2 & : \beta \leq v < \gamma' \\ 0 & : v \geq \gamma' \end{cases} \quad \dots \dots \dots \quad (9)$$

단, v 는 P_i 를 해결하는 데 소요된 시간, α 는 문제 해결에 허용된 하한 시간, $\gamma' = \gamma + \sigma$, β' 는 $(\alpha + \gamma')/2$ 를 의미한다.

1.4 난이도를 고려한 경우의 응답 정확도

난이도 문제를 해결하기 위해 전통적으로 사용해왔던 방법은 문항의 배점을 달리하는 것이었다. 즉, 평이한 문제의 배점을 1이라고 한다면 어려운 문제의 배점은 2, 지극히 어려운 문제의 배점은 3으로 부여하는 방식이다. 이 방식은 실제 대학 수학 능력 시험과 같은 매우 중요한 시험에서 실제 사용하고 있는 방식이다. 그러나, 이 방식은 학습자의 응답 성향과 관계없이 출제자의 사전 의도에 따라 미리 결정되는 단점이 있기 때문에 난이도라는 개념 자체에 존재하는 퍼지 속성을 반영하기에는 미흡한 점이 많다. 따라서, 본 연구에서는 퍼지 이론의 팽창 기법을 이용하여 난이도가 높은 문항에 가중치를 부여하는 방법을 제안한다. 이러한 방법으로 난이도가 고려된 경우의 응답 정확도 산출식 DCOR을 식 (10)에 나타냈다.

단, h 는 가중치 인자로 ‘쉽다’인 경우는 1, ‘보통이다’인 경우는 0.5, ‘중요하지 않다’인 경우는 0.25로 정의한다.

식 (10)에 의하면 다른 퍼지 속성 즉, 중요도 혹은 복잡도 등은 ‘보통이다’라는 퍼지 변수를 정상 변수로 본 것과는 달리 여기서는 ‘쉽다’라는 퍼지 변수를 정상 변수로 본 것이 차이가 있다.

2. 평가 함수와 지능적 평가

2.1 정규화

1절에서 구한 응답 정확도는 문항 당 m 개의 부문항으로 구성되어 있으므로 이를 응답 정확도는 최소 0, 최대 5의 범위 내에 있다. 따라서, 이들을 $[0, 1]$ 로 정규화하기 위해 다음 식 (11)~(14)와 같은 정규화를 수행한다.

단, m 은 각 문항에 포함된 부문항의 수를 의미한다.

또한, 문제 영역 전체에 대한 평가를 수행하기 위해 문제 영역 전체에 대한 정규화도식 (15)~식 (18)과 같이 수행한다.

단, n 은 문항의 수, m 은 각 문항에 포함된 부문항의 수를 의미한다.

2.2 평가 함수

이와 같이 얻어진 응답 정확도와 정규화 값을 통해 학습의 성취도를 평가하기 위해서 다음과 같은 퍼지 규속 함수를 정의한다.

‘매우 잘했다’ = x^2 ; $x=1$ 인 경우

‘잘했다’ = x ; $0 < x < 1$ 인 경우

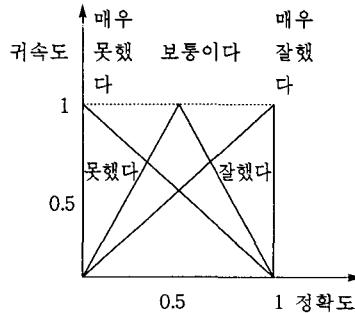
‘보통이다’ = $2x$; $0 < x \leq 0.5$ 인 경우

$$= -2x+2 \quad ; \quad 0.5 \leq x < 1 \text{인 경우}$$

$$\text{‘못했다’} = -x + 1 \quad ; \quad 0 < x < 1 \text{인 경우}$$

$$= (-x+1)^2$$

〈그림 3〉은 이들 폐지 귀속 함수를 그림으로 나타낸 것이다.



〈그림 3〉 평가를 위한 5개의 퍼지 귀속 함수

2.3 지능적 평가

각 문항은 먼저 식 (5), (7), (8), (9)에 의해 응답 정확도가 계산된 후, 식 (15)~(18)에 의해 정규화된다. 그러면 최종적으로 얻은 정규화된 응답 정확도는 앞의 퍼지 귀속 함수들 중의 하나에 의해 언어적으로 평가된다. 그런데, 정규화된 응답 정확도는 5개의 퍼지 귀속 함수 중 두 개 이상의 함수로 귀속될 가능성이 있다. 이와 같은 경우는 그 중 귀속도가 가장 큰 함수에 의해 적용되고 그에 해당되는 퍼지 언어 변수로 평가된다. 이를 수행하는 절차를 식 (19)~(22)에 나타냈다.

단, $FUZSET(MAX(P))$ 은 문항 P 의 정확도 $COR(P)$ 의 정규화 값인 $NORM(COR(P))$ 가 복수개의 퍼지 집합에 귀속되어 있을 때, 그 중 최대 귀속도를 보이는 퍼지 집합을 의미한다.

$$EV AL(P) = \bigcup_{i=1}^n \{P_i, NORM(ICOR(P_i))\} \\ = \bigcup_{i=1}^n \{P_i, FUZSET(\max_{j=1}^k NORM(ICOR(P_j)))\} \quad \dots \dots \dots (20)$$

$$\begin{aligned}
 EVAL(P) &= \bigcup_{i=1}^n \{P_i, NORM(CCOR(P_i))\} \\
 &= \bigcup_{i=1}^n \{P_i, FUZSET(\max(NORM(CCOR(P_i))))\} \dots\dots\dots(21)
 \end{aligned}$$

$$\begin{aligned}
 EVAL(P) &= \bigcup_{i=1}^n \{P_i, NORM(DCOR(P_i))\} \\
 &= \bigcup_{i=1}^n \{P_i, FUZSET(\max(NORM(DCOR(P_i))))\} \dots\dots\dots(22)
 \end{aligned}$$

IV. 모의 실험과 결과의 토론

1. 모의 실험 대상

본 연구에서 제안한 폐지 환경 하에서의 학습자의 응답에 대한 학습 성취도 평가를 수행하기 위해 다음과 같은 평가 문항을 만들었다. 이는 현직 초등학교 4학년 교사에 의해 작성된 것으로 복잡도, 중요도, 난이도 등이 문항에 내포된 것이다.

각 문항은 유형별로 5개의 문항과 각 문항에 부문항으로 5개의 문항으로 구성되어 있으므로 총 문항의 수는 25 문항이다. 문항 1에 해당되는 5개의 부문항은 모두 덧셈과 관련된 문항이다. 문항 2에 해당되는 5개의 부문항은 뺄셈에 관련된 문항이다. 문항 3에는 곱셈과 관련된 부문항들이 5개 있고, 문항 4에는 곱셈과 나눗셈이 혼합된 5개의 부문항이 있다. 마지막 문항 5에는 4칙 연산이 팔호와 혼합되어 구성되어 있는 5개의 부문항이 있다.

다음 〈표 1〉은 학습자에게 제시된 평가 문항이다.

〈표 1〉 학습자에게 제시된 평가 문항

부문항번호 문항번호	1	2	3	4	5
1	$3+5+7$	$13+5+7$	$13+15+7$	$130+5+7$	$130+105+7$
2	$8-3-5$	$18-3-5$	$18-13-5$	$180-13-5$	$18-13-51$
3	$3\times 5\times 7$	$13\times 5\times 7$	$13\times 15\times 7$	$130\times 5\times 7$	$130\times 105\times 7$
4	$16\div 4\div 2$	$160\div 16\div 2$	$47\div 3$	$5\times(12\div 6)$	$74\times 3\div 2$
5	$(8-5)\times(12\div 4)$	$13+20\div 5-10$	$17\times 4-64\div 16$	$15\div 3\times 7+9$	$7\times 8\div 28\times 3\div 2$

또한, 각 문항에는 다음과 같은 퍼지 요소가 내포되어 있는데, 이들은 문항의 복잡성을 나타내는 복잡도와 문항의 중요성을 나타내는 중요도와 문제 해결이 얼마나 어려운가를 나타내는 난이도로 구성되어 있다.

다음 <표 2>는 각 문항에 부여된 퍼지 속성이다.

<표 2> 각 문항에 부여된 퍼지 속성

문항	1			2			3			4			5		
	복잡	중요	난이												
1	S	NI	E	S	M	E	S	M	E	S	M	E	S	M	E
2	S	NI	E	S	M	E	S	M	E	S	M	E	S	M	E
3	S	NI	M	M	M	M	M	M	M	M	M	M	C	M	D
4	S	NI	M	M	M	M	S	M	M	M	I	M	C	M	D
5	S	I	D	C	M	D	C	M	D	C	M	D	C	M	D

단, ‘복잡’은 복잡도를, S는 ‘단순’, M은 ‘보통’, C는 ‘복잡’을 의미하고, ‘중요’는 중요도를, NI는 ‘중요하지 않음’, M은 ‘보통’, I는 ‘중요함’을 의미하고, ‘난이’는 난이도를, E는 ‘용이’, M은 ‘보통’, D는 ‘어려움’을 의미한다.

또한, <표 3>은 5개의 문항과 각각 5개의 부문항을 해결하는 데 허용된 하한 시간과 상한 시간이다.

<표 3> 문항 해결에 허용된 하한 시간과 상한 시간

단위 : 초, [하한시간, 상한시간]

문항	부문항				
	1	2	3	4	5
1	[1, 4]	[1, 4]	[1, 4]	[1, 4]	[1, 4]
2	[1, 5]	[1, 5]	[1, 5]	[1, 5]	[1, 6]
3	[3, 10]	[3, 10]	[3, 10]	[3, 15]	[3, 20]
4	[5, 10]	[5, 10]	[7, 15]	[7, 15]	[10, 15]
5	[10, 20]	[10, 20]	[10, 20]	[10, 20]	[20, 25]

2. 학습자의 응답 결과

〈표 1〉에 제시된 문항을 경기도 고양시의 한 초등학교 4학년 학생 10명에게 출제하여 그들의 응답 결과를 다음과 같은 〈표 4〉로 요약했다. 응답 결과에는 응답 결과의 정답 여부와 해당 문항을 푸는 데 소요된 시간을 같이 조사했다.

〈표 4〉 초등학교 4학년 학생 10명의 응답 결과표(S_1 (학생 1)~ S_{10} (학생 10))

소요시간 단위: 초

〈 S_1 〉

문항	부문항		1		2		3		4		5	
	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요
	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간
1	○	2	○	2	○	3	○	3	○	3		
2	○	2	○	2	○	4	○	4	○	5		
3	○	3	○	5	○	11	×	15	○	18		
4	○	9	×	8	○	13	○	12	○	16		
5	○	15	○	16	×	15	○	19	○	24		

〈 S_2 〉

문항	부문항		1		2		3		4		5	
	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요
	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간
1	○	2	○	1	○	2	○	2	×	3		
2	○	2	○	1	○	3	○	3	○	5		
3	○	5	○	7	○	7	×	9	×	11		
4	×	9	○	9	○	8	○	11	×	14		
5	○	21	○	21	○	21	○	29	○	29		

〈 S_3 〉

문항	부문항		1		2		3		4		5	
	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요
	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간
1	○	2	○	2	○	2	○	2	○	2		
2	○	1	○	1	○	3	○	5	×	6		
3	○	3	○	7	○	12	×	15	○	18		
4	○	14	○	13	○	12	○	12	○	12		
5	×	20	○	20	○	19	○	18	○	17		

〈 S_4 〉

문항	부문항		1		2		3		4		5	
	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요	정답	소요
	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간	여부	시간
1	○	2	○	2	○	2	○	2	×	2	○	3
2	○	2	○	3	○	2	○	2	×	5		
3	○	3	×	3	○	9	×	9	○	11		
4	○	9	○	9	○	8	×	9	○	9		
5	○	15	○	19	×	14	○	14	○	15		

$\langle S_5 \rangle$

부문항		1		2		3		4		5	
문항	부문항	정답	소요								
	여부시간										
1	O	2	O	2	O	2	O	3	O	3	
2	O	3	O	1	O	3	O	3	O	5	
3	O	3	X	3	X	5	O	7	O	7	
4	O	9	X	10	O	9	O	9	O	9	
5	X	15	O	15	O	15	O	15	O	17	

 $\langle S_6 \rangle$

부문항		1		2		3		4		5	
문항	부문항	정답	소요								
	여부시간										
1	O	2	O	1	O	2	O	2	O	2	
2	O	2	O	3	O	3	X	4	O	7	
3	X	3	O	7	O	9	O	12	O	15	
4	O	8	O	7	O	8	O	8	O	11	
5	O	16	O	18	X	19	O	19	O	24	

 $\langle S_7 \rangle$

부문항		1		2		3		4		5	
문항	부문항	정답	소요								
	여부시간										
1	O	2	O	2	X	2	O	2	O	2	
2	X	2	O	2	O	5	O	4	O	7	
3	O	4	O	5	O	5	O	10	O	12	
4	O	9	O	9	O	9	O	11	X	11	
5	O	15	X	15	O	17	O	19	O	20	

 $\langle S_8 \rangle$

부문항		1		2		3		4		5	
문항	부문항	정답	소요								
	여부시간										
1	O	2	O	2	O	3	O	3	O	2	
2	O	2	O	3	O	5	O	5	O	8	
3	O	3	X	4	X	7	O	10	O	14	
4	O	9	O	11	O	11	O	9	O	12	
5	X	15	O	18	O	21	O	21	O	21	

 $\langle S_9 \rangle$

부문항		1		2		3		4		5	
문항	부문항	정답	소요								
	여부시간										
1	O	2	O	3	O	2	O	2	X	2	
2	O	1	O	3	X	4	O	4	O	7	
3	O	3	O	3	O	9	O	9	O	13	
4	O	11	O	10	X	9	O	8	O	9	
5	O	15	O	26	O	20	O	21	O	21	

 $\langle S_{10} \rangle$

부문항		1		2		3		4		5	
문항	부문항	정답	소요								
	여부시간										
1	O	2	O	2	O	2	O	2	O	2	
2	O	2	O	2	O	3	O	5	X	6	
3	O	4	X	5	O	7	O	8	O	10	
4	O	10	X	10	O	10	O	9	X	9	
5	O	15	O	15	O	20	O	24	O	25	

3. 지능적 평가

3.1 문제 해결에 허용된 시간만 고려하는 경우

주어진 식 (5)에 의하여 테스트에 응한 학생들의 응답 결과에 대한 응답 정확도를 산출한다. 단, 지면 관계 상 10명의 학생 중 S_1 (학생 1)과 S_6 (학생 6)의 두명의 학생에 대해서만 산출하기로 한다.

〈표 5〉 응답 정확도(S_1 , S_6)

$\langle S_1 \rangle$						$\langle S_6 \rangle$					
부문 항 문 항	1	2	3	4	5	부문 항 문 항	1	2	3	4	5
1	0.778	0.778	0.222	0.222	0.222	1	0.778	1	0.778	0.778	0.778
2	0.875	0.875	0.125	0.125	0.08	2	0.875	0.5	0.5	0.125	0
3	1	0.837	0	0	0.028	3	1	0.367	0.041	0.125	0.173
4	0.08	0.32	0.125	0.281	0	4	0.32	0.68	0.969	0.969	0.92
5	0.5	0.32	0.5	0.02	0.08	5	0.32	0.08	0.02	0.02	0.08

이 두 학생에 대한 응답 정확도를 문제 영역 P에 대해 리스트로 표현하면 다음과 같다.

$$COR(P(S_1)) = \{(P_1, 2.222), (P_2, 2.08), (P_3, 1.865), (P_4, 0.726), (P_5, 0.92)\}$$

$$COR(P(S_6)) = \{(P_1, 4.112), (P_2, 1.875), (P_3, 0.706), (P_4, 3.858), (P_5, 0.5)\}$$

〈표 6〉 제한 시간만 고려한 경우의 평가

$\langle S_1 \rangle$						$\langle S_2 \rangle$						
정확도	정규화값	평가결과	정확도	정규화값	평가결과	정확도	정규화값	평가결과	정확도	정규화값	평가결과	
1	2.222	0.444	7.813	0.315	보통이다	'못했다'	1	4.112	0.8224	잘했다	11.051	0.442
2	2.08	0.416			보통이다		2	1.875	0.375	보통이다		
3	1.865	0.373			보통이다		3	0.706	0.1412	못했다		
4	0.726	0.145			못했다		4	3.858	0.7716	잘했다		
5	0.92	0.184			못했다		5	0.5	0.1	못했다		

이렇게 얻어진 결과를 이용하여 지능적 평가를 수행하면 〈표 6〉과 같다. 〈표 6〉에서 앞부분의 정확도, 정규화값, 평가 결과는 개별 문항에 대한 것이고, 뒷부분의 것은 문제 영역 전체에 대한 것이다.

3.2 중요도만을 고려하는 경우

주어진 식 (7)에 의하여 테스트에 응한 학생들의 응답 결과에 대하여 중요도를 고려한 경우의 응답 정확도를 산출한다. 역시, S_1 과 S_6 에 대해서만 수행한다.

〈표 7〉 중요도를 고려한 경우의 응답 정확도(S_1 , S_6)

		$\langle S_1 \rangle$					$\langle S_6 \rangle$						
		1	2	3	4	5			1	2	3	4	5
문항	부문항	1	0.778	0.222	0.222	0.222			1	0.778	0.778	0.778	0.778
		1	0.6053	0.778	0.222	0.222			1	0.778	0.778	0.778	0.778
1	1	0.7656	0.875	0.125	0.125	0.08			2	0.5	0.5	0.125	0
2	2	1	0.837	0	0	0.028			3	0.367	0.041	0.125	0.173
3	3	0.0064	0.32	0.125	0.5301	0			4	0.68	0.969	0.9844	0.92
4	4	0.7071	0.32	0.5	0.02	0.08			5	0.08	0.02	0.02	0.08
5	5												

이 두 학생에 대한 응답 정확도를 문제 영역 P에 대해 리스트로 표현하면 다음과 같다.

$$ICOR(P(S_1))$$

$$= \{(P_1, 2.0493), (P_2, 1.9706), (P_3, 1.865), (P_4, 0.6615), (P_5, 1.1271)\}$$

$$ICOR(P(S_6))$$

$$= \{(P_1, 3.9393), (P_2, 1.7656), (P_3, 0.706), (P_4, 3.6558), (P_5, 0.7457)\}$$

이렇게 얻어진 결과를 이용하여 지능적 평가를 수행하면 다음 〈표 8〉과 같다.

〈표 8〉 중요도를 고려한 경우의 평가

$\langle S_1 \rangle$					$\langle S_6 \rangle$					
	정확도	정규화값	평가결과		정확도	정규화값	평가결과	정확도	정규화값	평가결과
1	2.0493	0.4099	보통이다	7.6735	0.3069	'못했다'		10.8124	0.4325	'보통이다'
2	1.9706	0.3941	보통이다							
3	1.865	0.373	보통이다							
4	0.6615	0.1323	못했다							
5	0.1271	0.2254	못했다							

3.3 복잡도만을 고려하는 경우

주어진 식 (8)에 의하여 테스트에 응한 학생들의 응답 결과에 대하여 복잡도를 고려한 경우의 응답 정확도를 산출한다. 역시, S_1 과 S_6 에 대해서만 수행한다.

이와 같은 경우의 응답 정확도를 산출하기 위해 각 문항에 대해 응답한 10명의 학생들이 해당 문항을 푸는 데 소요한 시간의 표준편차를 구한 것이 〈표 9〉이다.

〈표 9〉 응답 정확도(S_1 , S_6)

문항	부문항	1	2	3	4	5
		1	2	3	4	5
1	0	0.54	0.4	0.46	0.49	
2	0.54	0.83	0.92	0.94	1.04	
3	0.66	1.58	2.21	2.62	3.3	
4	1.62	1.56	1.63	1.48	2.27	
5	2.18	2.1	2.51	4.09	4.08	

〈표 9〉로부터 최초 주어진 허용 시간의 하한과 상한을 조정한 시간 간격표를 〈표 10〉에 나타냈다.

〈표 10〉 조정된 하한 시간과 상한 시간

단위 : 초, [하한시간, 상한시간]

부문항 문항	1	2	3	4	5
1	[1, 4]	[1, 3.46]	[1, 3.6]	[1, 3.54]	[1, 3.51]
2	[1, 4.46]	[1, 4.17]	[1, 4.08]	[1, 4.06]	[1, 4.96]
3	[3, 9.34]	[3, 10]	[3, 10]	[3, 15]	[3, 23.3]
4	[5, 8.38]	[5, 10]	[7, 13.37]	[7, 15]	[10, 17.27]
5	[10, 17.82]	[10, 22.1]	[10, 22.51]	[10, 24.09]	[20, 29.08]

〈표 10〉을 통해 조정된 응답 시간의 귀속도를 이용하여 산출된 응답 정확도는 다음 〈표 11〉과 같다.

〈표 11〉 복잡도를 고려한 경우의 응답 정확도(S_1 , S_6)

$\langle S_1 \rangle$						$\langle S_6 \rangle$					
부문항 문항	1	2	3	4	5	부문항 문항	1	2	3	4	5
1	0.778	0.67	0.11	0.09	0.083	1	0.778	1	0.704	0.69	0.683
2	0.833	0.8	0.001	0	0	2	0.833	0.27	0.246	0	0
3	1	0.837	0	0	0.14	3	1	0.367	0.041	0.125	0.334
4	0	0.32	0.007	0.281	0.06	4	0.025	0.68	0.951	0.969	0.96
5	0.26	0.51	0.68	0.261	0.61	5	0.11	0.23	0.16	0.261	0.612

이 두 학생에 대한 응답 정확도를 문제 영역 P 에 대해 리스트로 표현하면 다음과 같다.

$$CCOR(P(S_1)) = \{(P_1, 1.731), (P_2, 1.634), (P_3, 1.977), (P_4, 0.348), (P_5, 1.641)\}$$

$$CCOR(P(S_6)) = \{(P_1, 3.855), (P_2, 1.349), (P_3, 0.867), (P_4, 3.585), (P_5, 1.213)\}$$

이렇게 얻어진 결과를 이용하여 지능적 평가를 수행하면 다음 〈표 12〉와 같다.

〈표 12〉 중요도를 고려한 경우의 평가

$\langle S_1 \rangle$						$\langle S_2 \rangle$							
	정확도	정규화값	평가결과	정확도	정규화값	평가결과		정확도	정규화값	평가결과	정확도	정규화값	평가결과
1	1.731	0.3462	보통이다	7.331	0.293	'못했다'		10.869	0.435	'보통이다'			
2	1.634	0.3268	못했다										
3	1.977	0.3954	보통이다										
4	0.348	0.07	못했다										
5	1.641	0.3282	못했다										

3.4 난이도만을 고려하는 경우

주어진 식 (9)에 의하여 테스트에 응한 학생들의 응답 결과에 대한 응답 정확도를 산출한다. 역시, S_1 과 S_6 에 대해서만 수행한다.

〈표 13〉 응답 정확도(S_1 , S_6)

		$\langle S_1 \rangle$							$\langle S_6 \rangle$				
		1	2	3	4	5			1	2	3	4	5
부문	항						부문	항					
1	1	0.778	0.778	0.222	0.222	0.222	1	0.778	1	0.778	0.778	0.778	0.778
2	2	0.875	0.875	0.125	0.125	0.08	2	0.875	0.5	0.5	0.125	0	0
3	3	1	0.915	0	0	0.167	3	1	0.606	0.202	0.354	0.416	0.416
4	4	0.283	0.566	0.354	0.53	0	4	0.566	0.825	0.984	0.984	0.959	0.959
5	5	0.841	0.752	0.841	0.376	0.532	5	0.752	0.532	0.376	0.376	0.532	0.532

이 두 학생에 대한 응답 정확도를 문제 영역 P 에 대해 리스트로 표현하면 다음과 같다.

$$DCOR(P(S_1)) = \{(P_1, 2.222), (P_2, 2.08), (P_3, 2.082), (P_4, 1.167), (P_5, 2.501)\}$$

$$DCOR(P(S_6)) = \{(P_1, 4.112), (P_2, 1.875), (P_3, 1.578), (P_4, 4.318), (P_5, 2.192)\}$$

이렇게 얻어진 결과를 이용하여 지능적 평가를 수행하면 다음 〈표 14〉와 같다.

〈표 14〉 난이도만 고려한 경우의 평가

$\langle S_1 \rangle$						$\langle S_2 \rangle$							
	정확도	정규화값	평가결과	정확도	정규화값	평가결과		정확도	정규화값	평가결과	정확도	정규화값	평가결과
1	2.222	0.444	보통이다	10.052	0.402	'보통이다'	1	4.112	0.8224	잘했다	14.075	0.563	'보통이다'
2	2.08	0.416	보통이다				2	1.875	0.375	보통이다			
3	2.082	0.416	보통이다				3	1.578	0.316	못했다			
4	1.167	0.233	못했다				4	4.318	0.864	잘했다			
5	2.501	0.5	보통이다				5	2.192	0.438	보통이다			

4. 결과의 토론

본 장에서는 학습자에게 임의의 테스트 문항이 주어졌을 때 학습자의 응답 결과를 통해 최적의 평가 결과를 제시할 수 있는 기법을 제안했다. 이 때, 테스트 문항에 중요도, 복잡도, 난이도 등과 같은 폐지 속성이 부여되고 이와 같은 폐지 속성을 처리하기 위해 폐지 이론을 이용하여 각각 최적의 공식을 제공했다. 본 연구에서 제안된 방식은 전통적인 평가 방식보다 훨씬 융통성 있는 평가 결과를 생성할 수 있음을 알 수 있다. 다음 〈표 15〉에 S_1 , S_6 의 두 학생에 대한 평가 결과를 상호 비교했다.

〈표 15〉 전통적인 평가 결과와 제안된 방법의 평가 결과 비교

구분	산술적 점수	폐지 속성을 고려한 평가
전통적 방식	$S_1 : 88점$ $S_6 : 88점$?
제안된 방식	$S_1 : 0.315$ $S_6 : 0.442$	$S_1 : '못했다'$ $S_6 : '보통이다'$

〈표 15〉에 의하면 전통적 방식을 이용한 평가의 최종 목적은 산술적 점수를 산출하기 위함이다. 다만 도출한 점수를 이용하여 미리 정해진 기준에 의해 'A', 'B' 혹은 '수', '우' 등을 부여할 수 있을 뿐이다. 그러나, 본 연구에서 제안한 방법은 산술적 점수를 도출하는 것이 목표가 아니라 언어적 평가를 하는 것이 목표이기 때문에 폐지 속성을 고려한

결과 ‘잘했다’, ‘못했다’ 등의 평가를 할 수 있는 것이다. 따라서, 최근의 평가 경향에 부합되는 적응적인 평가 방식이라고 생각할 수 있다.

V. 결 론

본 연구에서는 전통적으로 사용해왔던 학습자의 학습 성취도 평가 방식 대신 다양한 폐지 환경 하에서의 새로운 지능적 평가 방식을 제안했다. 그리고, 현직 교사라는 영역 전문가에 의해 출제된 학습 성취도 평가 문항을 초등학교 4학년 학생 10명을 대상으로 테스트했다. 그 결과, 출제 문항에 내포된 복잡도, 중요도, 난이도와 같은 다양한 폐지 환경을 종합적으로 반영하여 인간 교사가 직관적으로 생각할 수 있는 평가 결과와 유사하게 평가할 수 있음을 확인했다. 이것은 90점 이상이면 ‘수’, 80점 이상이면 ‘우’ 등으로 고정적으로 점수를 분할하여 평가해왔던 기존의 방식을 지양하고, ‘잘했음’, ‘못했음’과 같이 언어적으로 평가하고자 하는 새로운 교육 평가 방식에 지극히 부합되는 방법으로 생각된다. 따라서, 막연하게 언어적 평가를 할 수밖에 없었던 현재의 교육 현장에 새로운 평가 준거를 제공한다는 데 큰 의의가 있다고 사료된다. 그리고, 학습 성취도 평가 결과를 통해 향후 보다 과학적으로 새로운 학습 전략을 수립하는 데 많은 도움을 줄 수 있을 것으로 생각된다. 또한, 이를 폐지 속성을 통합하기 위해서는 주어진 각 속성들에 대한 식을 하나로 결합시키는 간단한 작업만으로도 가능하다는 것도 알 수 있었다. 그러나, 본 연구는 거의 모든 산술적 수치를 학습자의 응답 시간에 의존했다는 단점을 여전히 갖고 있다.

향후로는, 학습자의 전반적인 응답 성향을 분석하여 응답 정확도를 산출할 수 있는 방법이 모색되어야 할 것이다.

參 考 文 獻

1. 박성익(1988), 컴퓨터 보조 교육 공학—방법·개발·적용—, 교육과학사.
2. 변창진 외 공저(1986), 교육 평가, 학지사.
3. 서대우·이세훈·윤경섭·왕창종(1993), “지능형 저작 시스템에서 언어 변수를 이용한 학습자 모듈의 설계,” ’93 한국정보과학회 가을 학술 발표 논문집, Vol. 20, pp. 1289-1292.
4. 신동희·원성현·최성혜·정환목(1994), “지능형 교수 시스템에서 학습자의 인지 상태와 평가 방법에 대한 폐지 적용,” ’94 한국퍼지시스템학회 추계학술대회 논문집, Vol. 4, No. 2, pp. 152-156.
5. 신동희·원성현·정환목(1996), “폐지 이론을 적용한 교육 평가 방법에 관한 연구,” 한국퍼지및지능시스템학회 논문지, 제 6권 1호.
6. 원성현·엄정국(1992), 기초 폐지 이론과 응용 폐지 시스템, 정보시대 출판부.
7. 이민숙·곽덕훈·원유현(1992), “학습자 수준과 문제 난이도에 따라 시간을 부여하는 학습자 모듈의 설계,” ’92 한국정보과학회 가을 학술 발표 논문집, Vol. 19, No. 2, pp. 1083-1086.
8. 이옥화(1994), “교육의 과제와 테크놀로지의 역할,” 정보과학회지, Vol. 12, No. 6, pp. 16-29.
9. A. Barr and E. A. Feigenbaum(1981), *The Handbook of Artificial Intelligence*, Vol. 2, pp. 225-294, William Kaufmann Inc..
10. G. Kearsley(1987), *Artificial Intelligence and Instruction : Application and Method*, Addison Wesley Publishing Company.
11. Chiu-Keung Law(1996), “Using Fuzzy Numbers in Educational Grading System,” *Fuzzy Sets and Systems*, Vol. 83, No. 3, pp. 311-324.
12. Toshio Okamoto(1994), “The Current Situations and Future Directions of Intelligent CAI Research/Development,” IEICE Trans. on Information & System, Vol. E77-D, No. 1, pp. 9-18, Jan..
13. R. R. Yager(1981), “A Procedure for Ordering Fuzzy Subsets of the Unit Interval,” *Information Science*, Vol. 24, pp. 143-161.

14. L. A. Zadeh(1965), "Fuzzy Sets," *Information & Control*, Vol. 8, pp. 338-353.
15. L. A. Zadeh(1976), "A Fuzzy-Algorithmic Approach to the Definition of Complex or Imprecise Concepts," *International Journal of Man-Machine Studies*, Vol. 8, pp. 249-291.
16. 北垣 郁雄(1991), "ファジィ評価システムの開発研究," 電子情報通信學會論文誌, D-I, Vol. J74-D-I, No. 2, pp. 101-108.
17. 永岡 慶三・吳 亞棟(1991), "ファジィ変換に基づく學習應答時間マトリクスの解析," 電子情報通信學會論文誌, D-I, Vol. J74-D-I, No. 2, pp. 95-100.
18. 山下 元・清水 誠一(1993), "ファジィ理論を應用した評価システム ーファジィ推論による書藝の評價法ー," Tech. Report of IEICE ET93-21(1993-05), pp. 19-24.
19. 森田 龍彌・中島 和義・田村 正裕(1994), "評價函數にしなぶす荷重を考慮した多層ニューラルネットの教育學習法," 電子情報通信學會 信學技報, NC94-5, pp. 33-38.

Abstract

Design of Learning Achievement Evaluation Module of Intelligent Computer Assisted Instruction with Various Fuzzy Environment

Won, Sung-hyun

By decreasing in CPU price and development of computer assembling technology, personal computer take a good chance to accelerate its supply. Recently, as being introduced new computing technology so called multi media, learning assist system which is based on single media such as studying book, cassette tape, video tape, or something else is rapidly being replaced by new assist education system based on multi media in which it is operated by the personal computer.

In the computer assist education system, there is an evaluation module which appraise learner's study level into the next study strategy. At the view of this point, this part is very important. In this part, there are some factors like importance, complexity, or difficulty which commonly include fuzzy factors in our surrounding. But until now, we are still out of the level to handle the evaluation module adequately among the some studies.

In this study, we would like to suggest a new module that evaluate learning achievement of ICAI which have a variety of fuzzy environment. We combine independent fuzzy environment like importance, complexity, difficulty into making total evaluation of learner's achievement. By the result, with expressing by linguistic form, this study can provide the theoretical basis in which we will be able to carry out sentence toward evaluation among elementary school.