Journal of Korea Multimedia Society Vol. 10, No. 7, July 2007(pp. 912-922)

난이도 균일성을 고려한 유전자 알고리즘 기반 평가지 생성 시스템의 설계 및 구현

송봉기†, 우종호††

요 약

원격교육 시스템의 평가 시스템에서 평가의 공정성을 위하여 매 평가 시 평가지의 난이도를 일정하게 유지할 수 있는 방법이 요구된다. 본 논문에서는 유전자 알고리즘 기반의 평가지 생성 알고리즘을 제안한다. 평가지의 각 문항에 대한 난이도가 제출자에 의해서 지정되는 기존의 방법과는 달리 제안한 알고리즘에서는 각 문항의 난이도가 학생들의 평가 결과에 따라 적응적으로 조절되고, 평가지의 평균 난이도를 일정한 수준으로 유지할 수 있다. 제안한 알고리즘에서는 평가지에 동일한 문항이 중복으로 포함되는 것을 배제하고, 이전 평가의 결과를 반영하여 적응적으로 난이도가 조절될 수 있는 새로운 형태의 유전 연산자를 설계하고 구현한다. 그리고 모의실험을 통해 기존의 임의선택 방법과 모의 담금질 방법에 비해 균일한 난이도를 갖는 평가지가 생성될 수 있음을 보인다.

Design and Implementation of Genetic Test-Sheet-Generating Algorithm Considering Uniformity of Difficulty

Bong-Gi Song†, Chong-Ho Woo††

ABSTRACT

Evaluation of distance learning systems needs a method that maintains a consistent level of difficulty for each test. In this paper, we propose a new algorithm for test sheet generation based on genetic algorithm. Unlike the existing methods that difficulty of each test item is assigned by tutors, in the proposed method, that can be adjusted by the result of the previous tests and the average difficulty of test sheet can be consistently reserved. We propose the new genetic operators to prevent duplications of test items in a test sheet and apply the adjusted difficulty of each test item. The result of simulation shows that difficulty of the test sheet generated by proposed method can be more regular than the random method and the simulated annealing method.

Key words: Test Bank(문제은행), Test-Sheet-Generating Algorithm(평가지 생성 알고리즘), Genetic Algorithm(유전자 알고리즘)

1. 서 론

최근 정보통신 기술의 발달과 인터넷의 대중화에 따라 인터넷을 기반으로 하는 원격교육 시스템이 점차 일반화되고 있다. 원격교육 시스템은 웹 기반 교육(WBI: Web Based Instruction), 인공지능 기법을 적용한 지능형 교수 시스템(ITS: Intelligent Tutoring System) 등 다양하게 제안되고 연구되고 있

※ 교신저자(Corresponding Author) : 우종호, 주소 : 부산시 남구 대연3동 599-1(608-737), 전화 : 051)620-6881, FAX : 051)620-6450, E-mail : chwoo@pknu.ac.kr
접수일 : 2007년 1월 15일, 완료일 : 2007년 5월 22일

† 경성대학교 컴퓨터과학과
(E-mail : alphamir@ks.ac.kr)
†† 부경대학교 전자컴퓨터정보통신공학부
(E-mail : chwoo@pknu.ac.kr)

다[1,2]. 인터넷 기반의 원격교육 시스템은 다양한 형태의 교육 자료를 제공해줄 수 있고, 개인별 능력이나 학습속도에 따라 다양한 학습과정을 제공해 줄 수 있는 장점이 있다[3-5]. 이러한 원격교육 시스템에서 학생 개개인에게 맞는 교육을 제공하기 위해서 학생들의 학습이해도와 학습능력을 평가하는 일은 필수적이라고 할 수 있다. 정확한 학습평가는 학생들의 학습상태를 결정하기 위해서 중요한 기준이 된다. 컴퓨터 기반의 평가는 평가지가 동적으로 구성되고 멀티미디어를 활용하여 더 풍부한 평가요소들이 표현될 수 있으며, 학생들의 평가결과가 기록되고 분석될 수 있기 때문에 전통적인 종이와 연필로 하는 평가보다 더 효과적이고 유효하다고 밝혀지고 있다[4].

기존의 원격교육 시스템들에서 학생들의 학습 이해도에 대한 평가는 주로 문항(test item)들을 데이터베이스에 축적한 후 평가지(test sheet)의 크기에 따라 임의로 선택하여 생성하는 문제은행(test bank)시스템이 주로 이용되었고, 이에 대한 연구들도 활발히 이루어지고 있다[6]. 그러나 기존의 평가 시스템들에서는 교사가 그 자료나 문제들을 직접 관리하고 각 문항의 난이도를 주관적으로 판단하여 설정하기 때문에 학생들의 학습수준이나 환경적인 요소 등을 고려하기 어려웠다. 그 결과 생성된 평가지의 난이도가 평가 대상자 및 시험횟수에 따라 균일하지 못하고 정확한 평가를 하지 못하게 하는 주원인이 되었다.

본 논문에서는 유전자 알고리즘을 이용하여 학생들의 학습수준이나 환경적인 요소들을 고려하고, 평가 대상자 및 시험횟수에 따른 난이도의 편차를 줄이는 평가지 생성 알고리즘을 제안한다. 제안한 방법에서는 출제자가 문제은행에 문항을 등록할 때 각 문항의 초기 난이도를 설정하고, 매 평가가 이루어질 때마다 학생들의 평가결과에 따라 각 문항의 난이도를 적응적으로 조절한다. 이렇게 함으로써 문제은행의 각 문항들에 학생들이 실제로 체감하는 난이도가 반영될 수 있다.

제안한 방법에서 평가지의 문항들은 문제은행의 문항들을 대상으로 유전자 알고리즘을 적용하여 선택함으로써 생성된다. 문제은행에서 문항을 임의로 선택하여 염색체(chromosome)를 구성하고 교차를 수행하는 기존의 단순 유전자 알고리즘(SGA: Simple Genetic Algorithm)을 이 문제에 이용하면 평가지의 문항들이 중복으로 선택되어 하나의 평가지에 동일한 문항이 2개 이상 존재할 수 있다. 이 문제를 해결하기 위하여 제안한 방법은 염색체를 구성하는 염기(gene)들이 중복 발생하지 않는 새로운 교차연산자와 돌연변이연산자를 설계하고 구현한다. 또한 계산 속도를 높이고 유전자 알고리즘의 다양성을 유지하기 위하여 선택연산자로 룰렛 휠 선택과 엘리트 보존 기법을 병용하는 방법을 적용하였다.

모의실험을 통해 제안한 방법의 성능을 평가하였으며 그 결과 제안한 시스템이 임의선택 방법(Random Selection)을 사용하는 방법과 모의 담금질(Simulated Annealing)을 이용하는 방법에 비해 난이도를 더 균일하게 유지할 수 있음을 확인하였고, 제안한 유전자 알고리즘은 크기 10000의 문제은행에서 크기 100의 평가지 생성을 수행할 때 개체군의 크기가 30, 세대수 60 이상의 조건에서 더 효과적인 방법임을 확인하였다.

본 논문의 구성은 다음과 같다. 제2장에서 관련연구에 대해서 살펴보고, 제3장에서 본 연구에서 제안한 평가지 생성 알고리즘을 설명한다. 제4장에서는 제안한 평가지 생성 시스템을 구현하고, 제5장에서 모의실험을 통해 기존의 연구들과 비교·분석한다. 제6장에서 결론을 논의한다.

2. 관련 연구

황인수[6]는 모의 담금질 방법을 이용하여 주어진 제약조건을 만족하는 평가지 생성 알고리즘을 제안하였다. 이 연구에서는 문항의 유형, 문항의 난이도 등을 제약 조건으로 설정하고, 모의 담금질 방법을 이용하여 이러한 제약조건을 만족하는 평가지 생성을 수행하였다. 이 연구에서 사용된 에너지함수는 영역별로 출제하고자 하는 문항의 수(A)와 실제 출제된 문항의 수(B)의 함수로 표현된다. 이에 따라 첫 번째 에너지함수는 $|A - B|$를 사용하였으며, 두 번째 에너지함수는 A와 $|A - B|$의 비를 사용하였다[5]. 즉, 그 에너지함수는 다음 식 (1)과 식 (2)과 같다.

$$Energy_1 = |A - B| \tag{1}$$

$$Energy_2 = \frac{|A - B|}{A} \tag{2}$$

이 연구에서는 문제의 난이도를 제약조건으로

이용하였으나, 평가지의 생성 여부에 대한 연구만이 수행되었으며 난이도의 균일성에 대해서는 연구되지 않았다.

G. J. Hwang[3]은 평가의 질에 영향을 미치는 세 가지 요소를 정의하고 이를 기반으로 효과적인 평가를 수행할 수 있는 평가지 생성 알고리즘을 제안하였다. G. J. Hwang에 의해 정의된 평가의 질을 결정하는 세 가지 요소는 다음과 같다. 첫째, 각 문항의 난이도와 변별력 등의 정신적 요소에 대한 정확한 평가가 되었는가? 하는 것이다. 둘째, 문항들의 일관성 및 무결성을 유지하고 중복을 제거하는 문제은행의 관리기법이 적용되고 있는가? 하는 것이다. 마지막으로 다양한 요구 조건에 맞는 평가지 생성을 수행하는 알고리즘이 효율적인가? 하는 것이다.

G. J. Hwang은 연구를 통해 동적 프로그래밍(dynamic programming)을 이용한 평가지 생성 알고리즘을 제안하였다. G. J. Hwang의 방법에서 사용한 동적 프로그래밍의 식은 S_g와 E_g가 각각 평가지의 기대 난이도와 기대 분별력을 나타낼 때 식 (3)과 같다.

$$\min \sum_{i=1}^{N} (|S_i - S_g| + |S_i - S_g|) \tag{3}$$

이 때, S_i와 E_i는 Q_i는 i번째 수행단계에서 선택된 문항의 집합일 때 각각 다음 식 (4)과 같다.

$$S_i = \frac{S_{i-1} \times (i-1) + S[Q^i] \times 1}{i} \tag{4}$$
$$E_i = \frac{E_{i-1} \times (i-1) + E[Q^i] \times 1}{i}$$

G. J. Hwang의 방법에서는 평가지 생성을 수행할 때 개념과의 관계와 함께 난이도를 고려하고 있으나 난이도가 제출자에 의해 일방적으로 할당되기만 하고 학생들의 평가결과를 반영하지는 못한다. 또한 평가지 생성을 수행할 때 한 문항이 중복되어 선택될 수 있고 동적 프로그래밍의 특성상 입력된 데이터의 크기에 따라 지수 적으로 증가하기 때문에 대규모의 시스템에 적합하지 못하다.

G. J. Hwang[4] 등은 평가지 생성 시스템을 평가 유형에 따라 SLAT(Specified Length of Assesment Time)문제와 FNTI(Fixed Number of Test Items) 문제로 분류하고 이 두 유형에 대해 유전자 알고리즘을 이용한 평가지 생성 알고리즘을 제안하였다. SLAT는 평가시간을 제한한 경우에 개념들과의 관계와 변별력을 최적화하는 문제이고, FNTI는 고정된 수의 문항을 갖는 평가지에서 변별력을 최적화하는 문제이다. G. J. Hwang 등의 연구에서는 개념들과의 관계를 기준으로 평가지의 생성을 수행하는 CLFG(Concept Lower-bound First Genetic approach)와 평가지의 크기에 맞도록 문항을 선택한 후 더 적합한 문항으로 대체하는 FIFG(Feasible Item First Genetic approach)의 두 가지 유전자 알고리즘 기반 평가지 생성 알고리즘이 제안되었다. CLFG에서 염색체 부호화(chromosome encoding)는 선택된 문항을 1로 표현하고 선택되지 않은 것을 0으로 표현하는 이진 문자열 부호화를 사용하고, 선택연산자는 룰렛 휠 선택을 적용하였다.

문제은행을 구성하는 문항의 수가 n, i번째 문항의 변별력이 d_i, i번째 문항의 선택을 나타내는 부울 변수가 x_i, 평가시간이 하한보다 작은 경우에 선택된 문항의 평균 변별력에 대한 가중치가 α, 평가시간이 상한보다 큰 경우에 선택된 문항의 평균 변별력에 대한 가중치가 β, 페널티 점수가 R인 경우에서 제약조건의 근사치를 제한하는 평가함수 f는 다음 식 (5)과 같다.

$$f = \frac{\sum_{i=1}^{n} d_i x_i - \alpha - \beta - R}{\sum_{i=1}^{n} x_i} \tag{5}$$

이 알고리즘에서 교차연산자는 1점교차를 이용하며 돌연변이연산자는 전통적인 돌연변이를 이용하였다.

FIFG에서 선택연산자와 교차연산자는 CLFG방법과 동일하고, 염색체는 평가지의 각 문항에 대한 문제의 번호를 염기로 하여 부호화된다. 평가함수 f는 다음 식 (6)과 같다.

$$f = \sum_{i=1}^{n} d_{x_i} - R \tag{6}$$

돌연변이연산자는 변이율에 따라 문제은행에서 임의의 문항을 선택하여, 염색체의 임의의 위치에 있는 염기와 교체한다. 이 연구에서는 유전자 알고리즘 기반의 평가지 생성 알고리즘을 제안하였으나 문항의 중복과 난이도에 대한 고려가 배제되어 있다.

본 논문에서 제안하는 방법을 기존의 방법들과 비

교하면 다음 표 1과 같다. 제안한 유전자 알고리즘을 이용하여 평가지의 난이도를 균일하게 유지할 수 있고 문항의 중복을 배제하도록 유전 연산자들을 재설계한다.

3. 유전자 알고리즘 기반 평가지 생성 알고리즘

3.1 문제의 정의

원격교육 시스템에서 사용되는 평가 시스템은 다음 그림 1과 같이 구성된다. 문제은행에 저장된 문항을 평가지 생성 엔진에서 평가지의 크기에 맞춰 선택하고, 평가를 수행한다. 평가결과는 추론엔진에 전달되어 다음번의 평가를 위한 난이도의 조정과 전체 학습 진도를 조절하기 위하여 이용된다.

지능형 교수 시스템(ITS)의 평가 시스템에서 사용되는 평가지 선택 문제는 평가지의 문항들에 대해 개념들과의 관계 및 난이도에 의존적인 문제라고 할 수 있다. 즉, 평가지가 각 개념의 내용을 얼마나 반영하는 가에 따라 학생들이 취약한 개념을 찾는 기준이 되고, 평가지의 난이도에 따라 학생들의 학습에 대한 이해도를 평가할 수 있다. [3,7] 본 논문에서는 평가지를 생성할 때 난이도의 균일성을 유지하는 알고리즘을 개발하는 것이 목적이므로 개념들과의 관계에 대한 부분은 고려하지 않았다.

평가지의 문항에 대한 난이도 D는 교사가 초기에 설정하고, 학생들의 평가결과에 따라 조절된다. 즉, 교사가 설정한 난이도보다 더 낮은 정답률을 보이면 증가시키고, 더 높은 정답률을 보이면 감소시킨다. 초기에 설정된 난이도 D_{init}에서 유효한 인원 수 이상이 평가를 한 결과에서 정답률($D_{correct}$)이 D_{init}보다

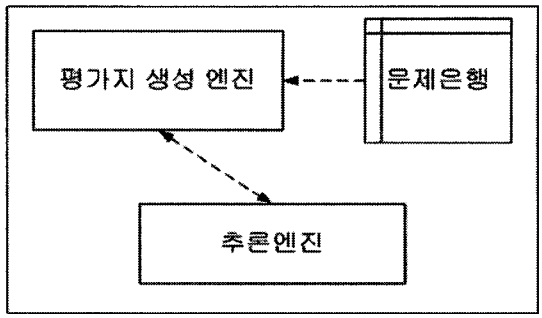


그림 1. 원격교육 시스템의 평가시스템

작으면, D_{init}에 $(100-D_{init})$와 $(1-D_{correct})$를 곱하여 더한다. 즉, 다음 식 (7)과 같이 나타낼 수 있다.

$$D = D_{init} + (100 - D_{init}) * (1 - D_{correct}) \quad (7)$$

정답률이 D_{init}보다 큰 경우에는 D_{init}에 $(1-D_{correct})$을 곱하여 뺀다. 이를 정리하면 다음 식(8)과 같이 나타낼 수 있다.

$$D = D_{init} * (1 - D_{correct}) \quad (8)$$

식 (7), (8)에서 정의된 난이도는 학습자가 평가를 하는 횟수가 증가함에 따라 학습자의 정·오답 비율에 따라 조정될 수 있음을 알 수 있다.

실제 평가지 생성 문제에서 문제은행에 있는 문제의 수가 n, 테스트를 구성하는 문제의 수가 m, 문제은행의 문제 집합 $Q = \{Q_1, Q_2, Q_3, \cdots, Q_n\}$, 평가지의 문제 집합 $S = \{S_1, S_2, S_3, \cdots, Q_m\}$인 조건에서 i번째 문제가 선택되는 것은 다음 식 (9)과 같고 m은 식 (10)과 같다.

$$x_i = \begin{cases} 0, & Q_i \notin S \\ 1, & Q_i \in S \end{cases} \quad (9)$$

표 1. 기존의 방법과 제안한 방법의 비교

비교항목		황인수의 연구[6]	Hwang의 연구[3]	Hwang 등의 연구[4]	본 논문의 방법
알고리즘		모의 담금질	동적 프로그래밍	유전자 알고리즘	유전자 알고리즘
난이도		고려함	고려함	고려되지 않음	고려함
각 문항의 난이도		교사가 설정	교사가 설정	교사가 설정	교사가 초기 설정, 평가결과에 따라 조정
유전연산자	선택연산자	-	-	룰렛 휠 기법	룰렛 휠 선택 기법과 엘리트 보존 기법
	교차연산자			1점 교차	새로운 연산자
	돌연변이 연산자			전통적 방법	전통적 방법

$$m = \sum_{i=1}^{n} x_i \tag{10}$$

이 때 생성된 평가지의 난이도 합계와 평균은 각각 식 (11), (12)과 같다.

$$D_{total} = \sum_{i=1}^{n} D_i x_i \tag{11}$$

$$D_{avg} = \frac{D_{total}}{m} \tag{12}$$

평가지 생성 문제는 평가지의 요구 난이도가 D_d인 경우에 다음 식(13)과 같은 최적화 문제로 정의될 수 있다.

$$\text{minimize } z = D_d - \frac{\sum_{i=1}^{n} D_i x_i}{m} \tag{13}$$

subject to

n >= m, n=0,1,2,…,n, m=0,1,2,…,m

$0 \le D_d \le 100,\ 0 \le D_i \le 100$

3.2 유전자 알고리즘의 설계

난이도의 균일성을 고려하여 평가지를 생성하는 문제는 NP-완전문제(NP-complete)이다[8]. J. H. Holland[9]에 의해 제안된 유전자 알고리즘은 생물의 유전 시스템을 모방하여 개체군에 선택, 교차, 돌연변이 연산들을 반복적으로 적용하여 근사 최적 해를 구하는 다점검색 기법으로 NP-완전문제에 대해 적용될 수 있다[10]. 본 논문에서는 유전자 알고리즘을 이용한 평가지 생성 알고리즘(Genetic Algorithm based Test-sheet Generating Algorithm, 이후 GATGA라 함)을 제안한다.

3.2.1. 염색체 부호화

GATGA에서 사용되는 염색체는 다음 그림 2와 같이 평가지 생성을 위하여 선택된 문제를 염기(Gene)로 하여 염색체를 구성한다. 그림 2에서 Q_1, Q_2, …, Q_n은 문제은행에서 선택된 평가지의 문항들이고, n은 평가지의 크기를 나타낸다.

문제은행의 문항들에서 평가지의 문항 수만큼 임의로 선택하여 염색체를 구성한다. 이 때, 동일한 문항의 중복은 허용하지 않는다.

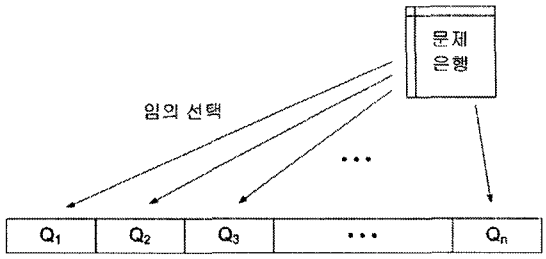


그림 2. 염색체의 부호화

3.2.2 평가 함수

GATGA에서 평가함수는 식 (13)의 조건을 고려하여 z가 최소화될 때 1로 되도록 설정하였다. 이는 다음 식 (14)과 같이 나타낼 수 있다.

$$f = \begin{cases} 1 - \dfrac{1}{\left| D_d - \sum_{i=1}^{n} \dfrac{D_i x_i}{m} \right|}, & \text{if } \left| D_d - \sum_{i=1}^{n} \dfrac{D_i x_i}{m} \right| \neq 0 \\ 1, & \text{if } \left| D_d - \sum_{i=1}^{n} \dfrac{D_i x_i}{m} \right| = 0 \end{cases} \tag{14}$$

평가 함수 f는 식 (14)의 D_d로 표현되는 기대 난이도와 생성된 평가지의 난이도 사이에 차가 발생하면 그 차의 비율만큼 1에서 감소시킴으로써 기대 난이도와의 오차가 클수록 적응도는 낮아지게 된다. 만약 기대 난이도와 평가지의 난이도가 일치하게 되면 가장 높은 적응도인 1을 갖는다.

3.2.3 선택연산자

유전자 알고리즘의 선택연산자는 개체군 내의 염색체들에 대해 각 염색체의 적응도의 분포에 따라 다음 세대에 생존하는 개체군을 확률적으로 결정한다. GATGA에서는 적응도에 비례하여 다음 세대에 생존할 확률이 결정되는 적응도 비례전략인 룰렛 휠 기법과 개체군 내에서 가장 적응도가 높은 염색체를 그대로 다음 세대로 남기는 엘리트 보존 기법을 혼합하여 사용하였다. 룰렛 휠 기법을 적용하여 우수한 적응도를 갖는 염색체 주변을 탐색함으로써 더 나은 적응도의 염색체를 찾을 확률을 높이고, 엘리트 보존 기법을 적용함으로써 계산 속도를 향상할 수 있다.

3.2.4 교차연산자

평가지 생성 문제에 대한 유전자 알고리즘 접근 방식에서 교차연산자는 염색체의 각 염기들이 중복

될 수 없는 특성을 반영하여야만 한다. 단순한 형태인 1점 교차 혹은 2점 교차를 사용하는 경우에는 교차연산의 결과로 염기의 중복이 발생할 수 있으므로 새로운 형태의 교차 연산자가 필요하다. GATGA에서는 계산 속도를 고려하여 단순한 형태인 1점 교차를 기반으로 하여 중복이 발생하지 않도록 교차연산자를 새롭게 설계한다. 제안한 교차연산자는 그림 3과 같다. A는 두 부모 염색체에 모두 포함된 문제 집합이며 교차연산과정에서 생략된다. 교차연산은 A를 제외한 B, C, D, E에 대해서 1점 교차를 수행한다.

부모 염색체 1, 2에서 서로 중복되는 염기를 선택하여 자식 염색체 1, 2의 앞부분으로 재배치하고, 나머지 영역에 대해서 1점 교차를 수행한다. 이 때 염색체의 염기의 수가 N_g이고, 중복되는 염기의 수가 N_d라면 교차 지점 P_c는 다음 식(15)을 만족해야만 한다.

$$N_d \le P_c \le N_g \tag{15}$$

3.2.5 돌연변이연산자

돌연변이연산자는 염색체의 염기를 일정한 확률로 변화시키는 조작으로 유전자 알고리즘에서는 이를 통해 지역 최적해(local optima)에서 탈출할 수 있다. 또한 돌연변이연산자는 유전자 알고리즘의 다양성을 확보하여 해의 질을 향상시킨다. GATGA에서는 전형적인 돌연변이연산자를 이용한다. 개체군에서 지정된 돌연변이 확률에 의해서 임의의 염색체를 선택하고, 염색체를 구성하는 염기 중 임의의 위치의 염기를 선택하여 문제은행의 문항으로 대체한다. 이 때, 중복을 방지하기 위하여 대체되기 위해 선택된 문항이 염색체 내에 존재하면 다시 선택하도록 한다.

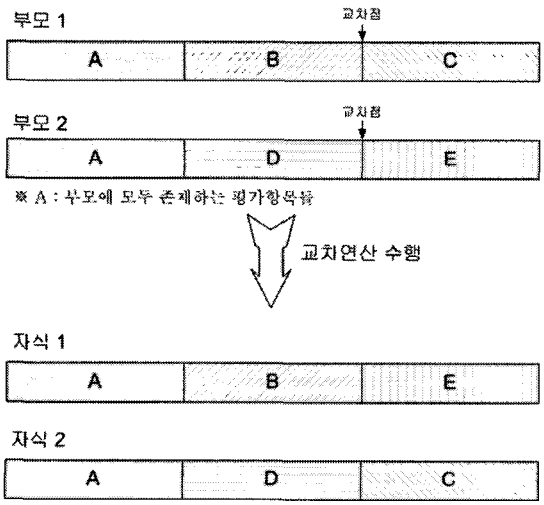


그림 3. 제안한 교차연산자

3.2.6 전체 유전자 알고리즘의 동작과정

새로 설계된 유전연산자들을 적용한 유전자 알고리즘의 동작은 그림 4와 같이 나타낼 수 있다. 먼저 염색체의 평가에 사용될 평가지의 난이도를 설정하고, 문제은행을 참고하여 염색체를 부호화한다. 염색체의 부호화에는 임의로 선택된 문제 번호와 문제번호에 부합되는 문제의 난이도 등이 문제은행으로부터 추출되어 이용된다. 부호화된 염색체를 이용하여 초기개체군을 생성한다. 개체군의 세대 수가 설정된 최대 세대 수에 도달할 때까지 선택연산자, 교차연산자, 돌연변이연산자를 적용하고 재생산을 반복한다. 최대 세대 수에 도달하게 되면 염색체를 역부호화를 하여 평가지의 생성을 수행한다. 이 때, 염색체에 있는 문제 번호에 준하여서 문제은행에서 문제를 추출하여 평가지가 생성된다. 필요에 따라 최대 세대 수를 조절함으로써 평가지의 질과 생성속도를 조절할

```
// Test-Sheet Generating Algorithm/
Set Expected Difficulty for Test Sheet
begin
  // GA //
  // generation : the current generation //
  // MAX_GENERATION : Termination condition
     of GA//
    Encode Chromosome;
    Generate Initial Population;
    while (generation != MAX_GENRATION)
      Select Operation;
      Crossover Operation;
      Mutation Operation;
      Reproduction;
    end while
    // Test-Sheet Generation //
    Decode Elite Chromosome;
    Generate Test-Sheet Using decoded Elite
    Chromosome;
end
```

그림 4. 제안한 유전자알고리즘의 동작

수 있다. 즉, 최대 세대 수가 클수록 더 나은 평가지의 질을 갖게 되나 생성속도가 느리고, 최대 세대 수가 작을수록 더 빠른 속도로 평가지 생성을 수행할 수 있으나 평가지의 난이도 균일성은 낮아지게 된다. 응용의 형태에 따라 최대 세대 수를 조절하여 이용할 수 있다.

4. 구 현

4.1 시스템의 구성

본 논문에서 제안한 평가지 생성 시스템은 다음 그림 5와 같이 유전자 알고리즘을 이용한 평가지 생성 모듈과 평가결과 관리 모듈로 구성된다. 평가지 생성 모듈의 역학은 문제은행으로부터 평가지에 포함될 문항을 추출하고 생성된 평가지가 평가 모듈에 전달되어 평가가 수행될 수 있도록 제어를 담당한다. 평가 모듈에 의해 수행된 평가결과는 난이도 관리 모듈에 전달되어 문제은행을 갱신하고 추론 엔진에 전달되어 학생들의 학습상황을 반영하게 된다.

4.2 시스템의 처리 과정

제안한 평가 시스템의 처리 과정은 그림 5에 나타난 것과 같이 문항추출, 평가지 전달, 평가결과 전달, 문제은행 갱신, 개인화된 학습전략 반영의 5단계를 갖는다. 각 단계에 대한 세부 항목은 다음과 같다. 첫 번째 단계인 문항추출 단계에서는 문제은행으로부터 문항을 추출하여 유전자 알고리즘 기반 평가지 생성 모듈에서 평가지의 생성을 수행한다. 이 단계에서 유전자 알고리즘을 이용하여 기대 난이도와 생성된 평가지의 난이도의 차이가 최소화되도록 문항이 추출되며, 문항의 중복이 없도록 교차연산자와 돌연변이연산자가 적용된다. 두 번째 단계인 평가지 전달 단계에서는 평가지 생성 모듈에서 생성된 평가지가

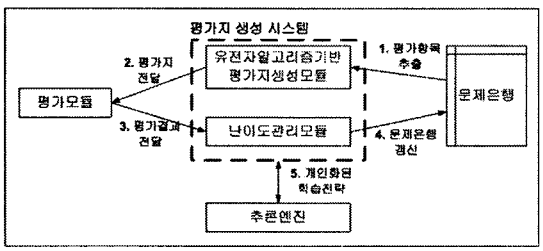


그림 5. 제안한 평가 시스템

평가 모듈에 전달되어 학생들이 평가를 수행하는 단계이다. 실시간 혹은 오프라인으로 학생들의 시험이 실시될 수 있다. 세 번째 단계인 평가결과 전달 단계에서 학생들의 평가결과가 난이도 조절을 위해 난이도 관리 모듈로 전달된다. 각 문항에 대한 학생들의 정답률과 오답률이 계산되며 이 계산된 결과가 난이도의 조정을 위하여 난이도 관리 모듈에 전단된다. 네 번째 단계인 문제은행 갱신 단계에서는 학생들의 평가결과를 반영하여 문제은행의 각 문항에 대한 난이도를 조절한다. 식 (3)과 식 (4)을 이용하여 정답률이 현재 난이도보다 높은 경우에는 난이도가 낮게 재설정하고, 정답률이 현재 난이도보다 낮은 경우에는 난이도를 높게 재설정한다. 마지막 단계인 개인화된 학습 전략 반영 단계에서 평가결과를 기준으로 추론 엔진에서 개인의 특성에 맞는 학습 전략을 평가 시스템에 전달받고 이를 적용한다. 재평가 수행여부, 기대 난이도 등이 추론 엔진으로부터 전달되어 설정된다. 각 단계는 학생의 평가결과를 난이도 조절에 적용할 수 있도록 유기적으로 연동된다.

5. 평가 및 고찰

5.1 평가조건

본 논문에서 제안한 GATGA에서 평가지의 난이도를 균일하게 유지하는 특성을 평가하기 위하여 기존의 원격교육 시스템에서 주로 이용되는 임의선택 방법, 모의 담금질 방법과 평가 함수의 적응도를 비교·평가하였다. 평가방법은 문제은행 크기의 변화에 따른 최적 적응도와 평가지의 문항 수 변화에 따른 최적 적응도를 비교·평가하였다. 또한, GATGA의 유전자 알고리즘 특성을 평가하기 위하여 세대수와 개체군의 크기 변화에 따른 최적 적응도의 변화를 모의실험 하였다. 모의실험에 적용한 기본 실험 조건은 다음 표 2와 같다. 표 2의 조건 중 유전자 알고리즘의 조건은 일반적으로 사용되는 조건을 적용하였다 [11].

5.2 문제은행의 크기에 따른 평가

문제은행의 크기에 따른 최적 적응도 비교함으로써 임의 선택 방법, 모의 담금질 방법, 제안한 방법의 특성을 평가하였다. 문제은행의 크기는 평가지 문항

표 2. 모의실험에 사용된 기본 조건

조건		값
문제은행의 크기		10000
평가지의 문항 수		100
목표 점수		70
유전자알고리즘	교차율	0.7
	변이율	0.3
	세대수	100
	개체군의 크기	30

수의 1.5배인 150, 2배인 200을 최소 조건으로 모의실험 하였고, 200 이상에서는 각 100의 변위로 평가지 문항 수의 10배인 1000까지 변하도록 설정하였다. 모의실험 결과는 다음 표 3과 같다. 표 3의 결과를 통해 제안한 방법이 임의 선택 방법에 비해서 15~20% 정도 우수한 결과로 평가지의 난이도를 유지함을 확인할 수 있고, 모의 담금질 방법에 비해서도 우수한 결과를 얻음을 확인할 수 있다. 문제은행의 크기가 200이상인 경우에는 0.98 이상의 적응도를 나타내었다. G. J. Hwang[3]의 연구에서는 동적 프로그래밍을 이용한 방법이 평균적으로 0.98-0.99의 결과를 보였다. 즉 제안한 방법이 동적 프로그래밍과도 거의 유사한 결과를 나타낼 수 있음을 확인할 수 있다. 유전자알고리즘의 특성상 세대수에 영향을 받을 수도 있기 때문에 이에 대한 평가는 5.4절에서 수행하였다.

표 3. 문제은행 크기에 따른 적응도 평가

문제은행의 크기	임의 선택 방법	모의 담금질 방법	제안한 방법
150	0.790	0.913	0.958
200	0.796	0.952	0.980
300	0.797	0.978	0.997
400	0.799	0.992	0.999
500	0.795	0.995	0.999
600	0.795	0.995	0.999
700	0.799	0.997	0.999
800	0.787	0.998	0.999
900	0.796	0.998	0.999
1000	0.800	0.998	0.999

5.3 평가지의 문항 수에 따른 평가

평가지의 문항 수에 따라서도 평가지의 난이도 균일성은 다른 특성을 나타낼 수 있기 때문에 평가지의 문항 수의 변화에 따른 최적 적응도도 평가하였다. 평가지의 문항 수에 따른 평가도 문제은행의 크기에 따른 평가와 동일하게 임의 선택 방법, 모의 담금질 방법, 제안한 방법에 대해 평가를 수행하였다. 평가지의 문항 수를 50 문항에서 100 문항까지 각 10 문항의 변위로 변화시키고, 100 문항 이상에 대해서는 변위를 크게 100 문항으로 설정하여 문항 수가 적은 경우와 큰 경우에 대하여 최적적응도를 평가하였다. 모의실험을 수행한 결과는 표 4와 같다. 표 4의 결과에서 제안한 방법이 임의 선택 방법, 모의 담금질 방법에 비해 효율적임을 확인할 수 있었다. 90문항 수 이상의 경우에는 모의 담금질 방법과 유사한 성능을 나타내나 문항 수가 50~80 문항일 경우에는 제안한 방법이 더 우수함을 확인할 수 있다.

5.4 세대 수와 개체군 크기에 따른 최적 적응도 평가

유전자 알고리즘은 동작 파라미터에 따라 그 성능에 영향을 받는다. 특히, 개체군의 크기와 세대 수의 변화는 최적 적응도의 수렴속도를 좌우하는 요소로서 매우 중요한 파라미터이다. 제안한 방법의 유전자 알고리즘 특성을 알기 위하여 개체군의 크기와 세대수의 변화에 따른 최적 적응도의 변화를 모의실험을 하였다. 개체군의 크기는 10에서 변위를 10으로 하여 100까지 변화하였고, 세대 수도 역시 10에서 변위를

표 4 .평가지의 문항 수에 따른 적응도 평가

평가지의 문항 수	임의 선택 방법	모의 담금질 방법	제안한 방법
50	0.79	0.99	1
60	0.79	0.99	1
70	0.79	0.99	1
80	0.79	0.99	1
90	0.79	0.99	0.99
100	0.79	0.99	0.99
200	0.80	0.97	0.98
300	0.80	0.93	0.94
400	0.80	0.92	0.92
500	0.79	0.90	0.90

10으로 하여 100까지 변화하였다. 모의실험 결과는 다음 그림 6과 같다. 그림 6의 결과를 통해 제안한 방법은 개체군의 크기가 30이상인 경우에 난이도를 일정하게 유지할 수 있고, 세대 수가 60 이상인 조건에서 최적 적응도를 나타냄을 알 수 있다. 즉, 표 2의 기본 조건을 적용한 경우에 제안한 방법을 적용할 때 개체군의 크기가 30, 세대 수 60 이상의 조건에서 난이도를 일정하게 유지함을 알 수 있다.

5.5 수행 시간에 따른 평가

유전자 알고리즘의 수행 시간은 세대 수와 개체군의 크기에 의존적인 특성이 있다. 표 2의 기본 조건을 적용한 상태에서 문항 수에 따른 수행시간을 모의실험하고 이를 임의 선택 방법과 비교하였다. 그 결과는 다음 표 5와 같다. 임의 선택 방법에 비해 제안한 방법이 더 긴 수행 시간을 요구함을 알 수 있다. 특히 유전자 알고리즘의 염색체가 문항 수에 비례하도록 설계되어 문제수가 많은 응용에서 더 긴 수행 시간이 요구된다.

5.6 고찰

모의실험의 결과를 통해 제안한 방법이 임의 선택 방법에 비해서는 더 나은 성능을 나타내고, 모의 담금질 방법이나 동적 프로그래밍 방법과는 유사한 성능 혹은 다소 우수한 성능을 나타냄을 확인할 수 있었다. 또한 유전자 알고리즘의 성능평가에서도 개체군의 크기 30, 세대 수 60 세대 정도의 적은 개체군의 크기와 적은 세대 수에서 최적 적응도에 수렴할 수 있음을 확인할 수 있었다. 수행시간 측면에서는 제안한 방법이 임의 선택 방법에 비해 수행 시간은 많이

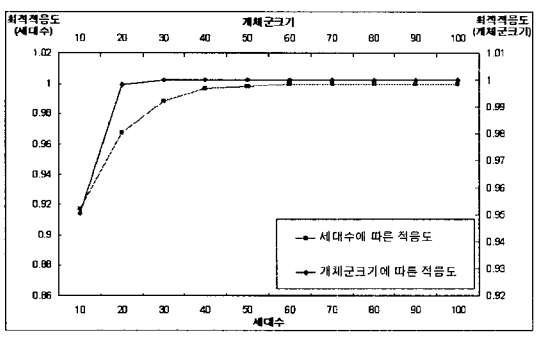


그림 6. 세대수와 개체군 크기에 따른 최적적응도 평가

표 5. 평가지의 문항 수에 따른 수행시간 평가

(단위: 초)

평가지의 문항 수	임의 선택 방법	제안한 방법
100	0.01	0.159
200	0.01	0.515
300	0.02	1.042
400	0.02	1.765
500	0.02	2.814

소요되나 일반적으로 사용될 수 있는 100-200문제 범위 내에서는 0.2초 정도로 수용할 수 있는 정도였으며, 평가지의 문항 수가 200문항 ~500문항 일 경우에는 임의 선택 방법에 비하여 수행시간이 많이 소요되나 약 2초 내외로 수용할 수 있을 정도이다.

모의실험 결과를 통해 본 논문에서 제안한 방법은 문제은행의 크기가 10000인 조건에서 문항 수 200개 이하의 크기를 갖는 평가지에 대해서 난이도를 균일하게 유지하고자 하는 응용에서 효과적인 방법이며, 개체군의 크기가 30이고 세대 수는 60 이상으로 함으로써 그 효과가 최대화될 수 있다. 임의 선택 방법에 비해 시간이 많이 소요되나 유전자 알고리즘의 특성상 개체군의 크기와 세대 수를 조절함으로써 줄일 수 있다. 문제은행의 크기에는 크게 영향을 받지 않으나 문제은행의 크기가 문항 수의 10배 이상이 되는 경우에 더 효율적인 방법이다.

6. 결 론

본 논문에서는 원격교육 시스템에서 더 정확하고 공정한 평가를 할 수 있도록 각 평가회수에 따른 평가지의 난이도 편차를 줄이는 방법으로 유전자 알고리즘 기반의 평가지 생성 알고리즘을 제안하였다. 제안한 방법에서는 학생들의 평가결과를 기반으로 각 문항의 난이도를 적응적으로 변화시킴으로써 각 문항의 난이도에 학생들의 체감 난이도가 적용될 수 있다. 또한 평가지 생성 시에 중복되는 문제가 선택되지 않도록 새로운 교차연산자를 설계하였으며, 선택연산자도 룰렛 휠 기법과 엘리트 보존 기법을 병용함으로써 전체 유전자 알고리즘의 수행성능도 향상시켰다. 모의실험을 통해 기존의 방법과 비교·분석한 결과 문항을 임의로 선택하여 평가지 생성을 수행

하는 임의 선택 방법의 경우 기대 난이도에 80% 수준으로 평가지가 생성되었으나, 모의 담금질 방법과 제안한 방법은 90% 이상의 수준을 보임을 확인하였다. 제안한 방법이 임의선택방법에 비해 10%이상 우수하게 평가지의 난이도를 균일하게 유지하고, 모의 담금질 방법에 비해서도 우수한 성능을 나타냄을 확인할 수 있었다. 또한 제안한 방법의 적용할 경우, 개체군의 크기가 30, 세대 수가 60 이상인 조건에서 그 효과가 최대화됨을 확인하였다. 본 논문에서는 평가지 생성 시 각 문항의 개념들과의 관계를 포함시키지 않았으나 향후 연구를 통해 더 다양한 조건 하에서 난이도의 균일성을 유지하는 평가지 생성 알고리즘을 개발하고, 추측에 의한 영향을 줄일 수 있는 방법을 개발하고자 한다.

참 고 문 헌

[1] S. Hopper, "Cooperative Learning and Computer-Based Instruction," *Educ. Technol. Res. Develop.*, Vol. 40, No. 3, pp. 21-28, 1992.

[2] G.J. Hwang, "A Tutoring Strategy Supporting System for Distance Learning on Computer Networks," *IEEE Trans. Educ.*, Vol. 41, No. 4, pp. 1-19, 1998.

[3] G.J. Hwang, "A Test-Sheet-Generating Algorithm for Multiple Assessment Requirements," *IEEE Trans. on Educ.*, Vol. 46, No. 3, pp. 329-337, 2003.

[4] G.J. Hwang, B.M.T. Lin, H.H. Tseng, and T.L. Lin, "On the Development of A Computer-Assisted Testing System With Genetic Test Sheet-Generating Approach," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. 35, No. 4, pp. 590- 594, 2005.

[5] B.A.A. Antao, A.J. Brodersen, J.R. Bourne, and J.R. Cantwell, "Building Intelligent Tutorial Systems for Teaching Simulation in Engineering Education," *IEEE Trans. Educ.*, Vol. 35, No. 1, pp. 50-56, 1992.

[6] 황인수, "사이버 문제은행에서 시뮬레이티드 어닐링을 이용한 2단계 문제세트 생성 휴리스틱," 한국경영과학회/대한산업공학회 춘계공동학술대회, 제18권, 제1호, pp. 155-164, 2001.

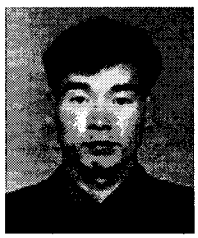
[7] 홍종기, 전우천, "수준별 평가를 위한 문제은행 시스템," 한국정보교육학회지, 제6권 제3호, 2003.

[8] H. Lange, R. Möller, and B. Neumann, "Avoiding Combinational Explosion in Automatic Test Generation: Reasoning about Measurements is the Key," *KI'96, Dresden*, pp. 213- 226, 1996.

[9] J.H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, 1975.

[10] T.Z. Zhang, Y.G. Fu, and R.M Shen, "Improve Question & Answer System by Applying Genetic Algorithm," *Proc. of the Third International conf. on Machine Learning and Cybertics*, Vol. 4, No. 1, pp. 2317-2321, 2004.

[11] J. Grefenstette, "Optimization of Control Parameters for Genetic Algorithms," *IEEE Trans. on System, Man and Cybernetics*, Vol. 16, No. 1, pp. 122-128, 1986.



송 봉 기

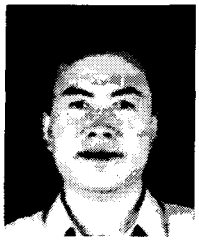
1991년~1997년 부경대학교 전자공학과 학사

1998년~2000년 부경대학교 정보시스템 협동과정 정보공학 석사

2000년~2005년 부경대학교 정보시스템 협동과정 정보공학 박사

2005년~현재 경성대학교 컴퓨터과학과 초빙외래교수

관심분야 : 유전자알고리즘, 네트워크, 마이크로프로세서, 임베디드시스템 등



우 종 호

1978년 2월 경북대학교 컴퓨터공학과(학사)
1981년 2월, 1986년2월 경북대학교 대학원 전자공학과 전산공학전공(석사, 박사)
1981년 3월~현재 부경대학교 전자컴퓨터정보통신공학부 교수
1987년 8월~1988년 8월 미국 RPI Dept of Computer Science, Visiting Scholar
2001년 8월~2002년 8월 미국 LSU at BR, Dept of ECE, Visiting Professor
관심분야 : 임베디드시스템, 센서네트워크, 바코드/RFID, e-learning