

---

# 효율적인 자기 주도적 학습 진단을 위한 문제 출제 알고리즘

김 은 정\*

Examination Questions Selection Algorithm for  
Efficient Self-Directed Learning diagnosis

Eun-Jung Kim\*

## 요 약

이러닝 문제은행기반 출제시스템에서 많은 학습자들은 난이도에 따른 자동 출제 문제들을 기반으로 스스로를 진단함으로써 자기 주도적으로 학습을 진행한다. 이러한 방식에서 학습자들의 효율적인 자기 주도적 학습 진행을 위해서는 정당한 방법으로 출제되는 문제가 무엇보다 중요하다. 본 논문에서는 효율적인 자기 진단 평가를 위해, 평가의 종류에 따라 난이도뿐만 아니라 학습 범위와 문제의 키워드를 함께 고려하는 새로운 방법의 문제 출제 알고리즘을 제시한다. 이는 기존의 난이도만을 고려한 자동 출제 방식에 비해 자기 주도적 학습자에게 보다 효과적인 학습 진단 방법을 제공할 수 있다.

## ABSTRACT

Many learners on E-learning databank based selection system making self-directed progress with learning by diagnosis oneself based on automatically selection questions using degrees of difficulty. This methods is most important to choose a questions using right a way for effective self-directed learning progress of learners. This paper present new question selection algorithms consider for degree of difficulty, scope of learning and keyword of questions according to examination type. This algorithm providers more effective learning diagnosis methods as compared with previous algorithm consider for only degrees of difficulty

## 키워드

자동 문제 출제 시스템, 문제은행 출제 시스템, 웹 기반 학습 시스템, 원격 교육 시스템

---

\* 부산가톨릭대학교 컴퓨터공학과

접수일자 2009. 02. 18  
심사완료일자 2009. 04. 20

## I. 서 론

자기 주도적 학습(Self-Directed Learning)은 학습자 스스로 학습 목표를 설정하고 학습 과정 및 학습 자원을 결정하여 학습을 수행하고 학습 결과를 스스로 평가하는 일련의 학습 과정을 말한다. 이러한 자기주도성은 e러닝에서 가장 중요한 요소 중 하나이며, 효율적인 자기 주도적 학습 진행을 위해서는 학습 결과를 스스로 평가하는 평가 문제와 문제를 출제하는 방식이 무엇보다 중요하다고 할 수 있다.

일반적으로 웹기반 학습 시스템에서 평가를 위해서는 문제 응행 방식을 많이 이용한다. 이러한 학습 시스템에서 보다 효과적인 평가를 위해서 문제 출제 방식에 대한 연구가 많이 있어 왔다[1-4]. 일반적으로 이러닝에서 평가를 위해 사용하는 출제 방식으로는 고정 출제나 무작위 출제 방식 또는 난이도에 따른 자동 출제 방식이 있다. 이 중에서 난이도를 이용한 무작위 출제 방법을 살펴보면, 이는 문제를 데이터베이스에 등록할 때 해당 문제의 난이도를 입력하고 평가 결과를 기반으로 지속적으로 난이도를 관리함으로써 자동 문제 출제에 이를 이용한다. 이 방법은 학습자들의 학습 능력을 고려하여 적절한 문제 출제가 가능하기 때문에 이러닝에서 주로 많이 이용되고 있으며, 본 논문에서도 이 방법을 사용한다.

이러닝의 자기 주도적 학습 진행이 효율적으로 진행되기 위해서는 여러 종류의 평가가 이루어져야 한다. 또한 각 종류의 평가가 갖는 의미가 최대의 효과를 얻기 위해서는 해당 평가 유형에 맞는 문제가 출제되는 것이 무엇보다 중요하다. 이에 본 논문에서는 자기 주도적 학습자들을 위한 웹기반 학습 시스템을 설계함에 있어서 평가를 위한 문제 출제 방법에 초점을 맞추고자 한다. 일반적으로 학습 평가 종류에는 각 단원별 평가와 과목별 평가, 그리고 틀린 문제에 대한 재학습을 한 후의 재평가가 있다. 첫째 특정 단원의 학습이 끝난 다음에 학습 결과를 진단하는 단원별 평가는 특정 단원에 속하는 문제에 대해서만 난이도를 고려하여 문제를 출제하면 된다. 둘째 특정 과목에 대한 전체 학습 후에 이루어지는 과목별 평가는 과목의 전체 영역에서 문제를 끌고루 출제해야 하는 특징이 있다. 따라서 문제의 난이도만을 고려하다 보면, 문제 출제가 특정 단원에 치우치거나 특정 단원에서는 한 문제도 출제되지 않는 경우가 발생할 수 있다. 따라서 예상 평균 점수에 맞는 난이도를 고려할 뿐만 아니라

라 전체 단원에서의 문제 분포를 함께 고려하는 새로운 난이도/영역별 알고리즘을 제시한다. 셋째 틀린 문제와 관련된 학습 영역을 재학습 한 다음에 이루어지는 재평가에서는 해당 문제와 관련이 있는 문제 위주의 출제가 이루어져야 하는 특징이 있다. 특정 문제와 관련이 있는 문제들은 특정 단원에 한정되어 있을 수도 있고, 전체 단원에 분포되어 있을 수 있다. 따라서 전체 단원에서 각 문제의 키워드를 고려하여 틀린 문제와 관련된 문제에 대해서만 난이도를 고려하여 출제하는 난이도/키워드 별 알고리즘을 제시한다. 이러한 평가 유형에 따른 서로 다른 출제 알고리즘의 적용은 자기주도적 학습을 진행하는 학습자들에게 보다 효율적인 진단 프로그램을 제공할 수 있다.

## II. 관련 연구

웹기반 교육에서의 평가 방법에 관한 많은 연구 중에서 문제 응행 방식을 이용한 무작위 문제 출제에 관한 관련 연구는 다음과 같다.

[3]에서는 학습자의 학습 진행 상태 및 학업 성취도에 따라 수준별로 제시된 문제를 풀게 한다. 각 수준별 문제는 학습의 진행 상태에서 무작위로 출제되는 문제를 풀고 다음 단계로 넘어가는 것이다. 이는 문제의 수준에 대한 평가가 출제자의 임의의 판단 기준에 의한 것이므로 객관적인 근거가 없고 각 단계별 수준에 대한 신뢰성이 떨어진다. [4]에서는 난이도에 따른 자동 문제 출제 시스템을 구현함에 있어서, 5단계의 난이도를 기반으로 예상 평균 점수와 출제 문제수에 따라 출제되는 문제들의 정답률의 평균이 예상평균점수가 되도록 자동 출제 한다. 여기서는 예상평균점수에 따른 정답률별 문제수 비율을 표준정규분포 방식에 의해 구하였다. 본 논문에서의 난이도별 문제 수 비율도 이 연구의 근거에 의하고 있다. 그러나 이러한 난이도별 자동 문제 출제 시스템은 문제 응행에 등록된 문제 중에서 난이도만을 고려하여 해당 비율의 문제들이 출제되기 때문에 난이도의 객관성이 입증되었다 하더라도 문제가 특정 영역에 편중되어 출제될 수 있기 때문에 문제 영역의 전반적인 평균 출제는 이루어지지 않는다. 따라서 기존의 난이도별 문제 출제 방식으로는 전체 영역이 고려되어야 하는 과목별 평가나 특정 문제와의 관련성을 고려해야 하는 재평가

와 같은 시험 방식에는 적합한 문제 출제 알고리즘이 될 수 없다. 즉 평가의 유형에 따라 난이도뿐만 아니라 전체 학습 영역과 문제의 키워드 등을 함께 고려할 수 있는 새로운 출제 알고리즘이 필요하다.

### III. 평가 유형별 문제 출제 알고리즘

#### 3.1 데이터베이스 구조

이전의 연구[5]에서 웹기반 학습 시스템을 설계한 바 있다. 온라인 학습 시스템에서 사용하는 데이터베이스는 전체 7개의 테이블로 구성되어져 있다. 이 중에서 본 논문에서는 문제 출제에 초점을 맞추고자 하므로 시험 문제를 관리하는 [과목별문제관리테이블]에 대한 구조만을 다루고자 한다. 표 1에서 각 문제는 문제가 속한 단원과 문제, 보기, 정답이 등록된다. 문제의 난이도는 ‘상’, ‘상중’, ‘중’, ‘중하’, ‘하’로 구분한다. 이러한 난이도는 정답률과 오답률에 기초하여 표 2와 같이 난이도 비율로 관리된다. 정답률과 오답률에 따른 난이도 비율의 계산 방법 및 문제의 동적 난이도 재조정에 대해서는 기존의 연구[7]에서 설명한 바 있다. 키워드1, 키워드2는 해당 문제가 어느 영역 또는 개념에 속하는지를 설명할 수 있는 대표되는 용어를 순서대로 저장하고 키워드3은 해당 문제와 연계되는 영역을 설명할 수 있는 용어를 저장한다.

표 1. ‘과목별문제관리’ 테이블의 구조

Table 1. Structure of ‘question’ table

필드명	데이터형식	비고
번호	int	각 문제를 식별하는 번호
단원	char	문제가 출제된 단원
문제	string	
보기1	char	
보기2	char	
보기3	char	
보기4	char	
정답	char	정답 표기(예: 보기3)
난이도	char	초기치 ‘중’
난이도비율	int	초기치는 50(백분율로 표기)
키워드1	string	문제에 대한 키워드 1
키워드2	string	문제에 대한 키워드 2
키워드3	string	문제와 연계되는 개념의 키워드
문제등록일	date	최초 문제 등록일
최종수정일	date	가장 최근의 난이도 조정일

표 2. 난이도별 비율

Table 2. Degrees of difficulty rate

난이도	난이도비율	비고
상	20% 이하	아주 어려운 문제
상중	21 ~ 40%	어려운 문제
중	41 ~ 60%	보통인 문제
중하	61 ~ 80%	쉬운 문제
하	81% 이상	아주 쉬운 문제

이는 틀린 문제에 대한 재학습 결과를 진단하는 재평가를 위한 문제 출제에 이용된다.

#### 3.2 평가 유형별 문제 출제 알고리즘

본 논문에서는 평가 유형을 다음의 세 가지로 나누어 관리한다. 먼저 각 단원마다 수행되는 단원별 평가와 특정 과목에 대해 수행되는 과목별 평가 그리고 틀린 문제에 대한 재학습을 한 후에 수행되는 재평가로 이루어진다.

##### 3.2.1 단원별 평가

각 단원에 대한 학습후에 학습자는 해당 단원에 속하는 문제를 대상으로 학습 결과를 평가하고 스스로 진단하여 다음 학습을 진행한다. 이에 본 논문에서는 특정 단원을 평가함에 있어서 학습자의 학습 능력을 고려하기 위해 문제의 난이도별 출제 방법을 적용한다. 평균 점수에 대한 난이도별 출제 문제 수 비율은 [4]에서 제시한 예상평균점수에 따른 정답률별 문제수 비율을 근거로 한다. 여기서 예상평균점수란 출제되는 문제들의 정답률의 평균이 예상평균점수가 되도록 한다는 것이다. [4]의 난이도별 자동 문제 출제 알고리즘에서는 표준정규 분포 방식[6]에 의해 예상 평균점수에 따른 정답률별 문제 수 비율을 구하였다. [4]에서 구한 예상평균점수에 따른 정답률별 문제수 비율을 요약하면 표 3과 같다. 표 3을 바탕으로 표 4는 총 20문제 출제에 대해 예상평균점수에 따른 난이도별 출제 문제수 비율을 계산한 결과이다. 표 4에서 예상 평균점수 50점에 맞추어 출제할 경우, 전체 20문제 출제에 ‘상’ 1문제, ‘상중’ 5문제, ‘중’ 8문제, ‘중하’ 5문제, ‘하’ 1문제가 출제 되므로, 난이도가 ‘상’인 문제는 전체 5%, ‘상중’은 전체 25%, ‘중’은 전체 40%, ‘중하’는 전체 25%, ‘하’는 전체 5%의 비율로 출제된다. 이에 대한 예상평균점수에 따른 단원별 자동 문제 출제 알고리즘은 그림 1과 같다.

표 3. 예상평균점수에 따른 정답률별 문제수비율(%)  
Table 3. question rate by correct-rate according to prediction average score

난이도(정답률)	상	상중	중	중하	하
90점			7	24	69
80점		2	14	34	50
70점		7	24	42	27
60점	2	14	34	36	14
50점	7	24	38	24	7
40점	14	36	34	14	2
30점	27	42	24	7	
20점	50	34	14	2	
10점	69	24	7		

표 4. 난이도별 출제 문제수 비율  
Table 4. examination question rate by Degrees of difficulty

출제 비율(출제 문제수)	예상평균점수	총 시험문제(20문제)				
		상	상중	중	중하	하
출제 비율(출제 문제수)	90점	0% (0)	0% (0)	5% 1.40(1)	25% 4.8(5)	70% 13.8(14)
	80점	0% (0)	0% 0.4(0)	15% 2.8(3)	35% 6.8(7)	50% 10(10)
	70점	0% (0)	10% 1.40(2)	25% 4.8(5)	40% 8.4(8)	25% 5.4(5)
	60점	0% 0.4(0)	15% 2.8(3)	35% 6.8(7)	35% 7.2(7)	15% 2.8(3)
	50점	5% 1.40(1)	25% 4.80(5)	40% 7.60(8)	25% 4.80(5)	5% 1.40(1)
	40점	15% 2.8(3)	35% 7.2(7)	35% 6.8(7)	15% 2.8(3)	0% 0.4(0)
	30점	25% 5.4(5)	40% 8.4(8)	25% 4.8(5)	10% 1.4(2)	0% (0)
	20점	50% 10(10)	35% 6.8(7)	15% 2.8(3)	0% 0.4(0)	0% (0)
	10점	70% 13.8(14)	25% 4.8(5)	5% 1.4(1)	0% (0)	0% (0)

```

function 문제출제(문제수, 예상평균점수)
    난이도별출제갯수(5,2) = "예상평균점수에 따른 '난이도' , '개수'를 초기화"
    for i = 0 to 4   ' 5단계의 난이도
        출제난이도 = 난이도별출제갯수(i,0)
        레코드셋 = select from 과목테이블 where 출제난이도
        출제문제수 = 난이도별출제갯수(i,1)
        count = 0
        do while(count < 출제문제수)
            난수(레코드번호) = int((total_count-1 + 1) * rnd() + 1 )
            해당 레코드 문제를 사용자 시험 데이터베이스에 출제
            count = count + 1
        loop
    end for
end function

```

그림 1 난이도별 문제출제 알고리즘

Fig. 1. question selection algorithm by Degrees of difficulty

### 3.2.2 과목별 평가

특정 과목에 대한 학습이 끝난 후, 과목 전체를 평가하기 위해서는 문제가 특정 단원에서 집중적으로 출제되거나 특정 단원이 출제되지 않는 경우가 없이 대체적으로 학습 전 영역에서 꿀고루 문제가 출제되어야 한다. 이에 전체 단원과 난이도가 함께 고려되는 난이도/영역 별 문제 출제 알고리즘을 적용한다.

먼저, 난이도에 대한 출제 문제수 비율은 표3을 바탕으로 계산한 표4에 따른다. 다음으로 해당 과목의 전체 단원 수를 고려하여 영역별 출제 문제수 비율을 정한다. 총 20문제를 출제함에 있어서 전체 단원 수에 대한 출제 문제 수 비율은 표 5와 같다.

### 표 5. 전체 단원수별 출제 문제수

Table 5. The number of selection question by all units

총 단원수	출제 문제수 - 총시험문제(20문제)									
	1장	2장	3장	4장	5장	6장	7장	8장	9장	10장
2장	10	10								
3장	6(+1)	6(+1)	6(+1)							
4장	5	5	5	5						
5장	4	4	4	4	4					
6장	3(+1)	3(+1)	3(+1)	3(+1)	3(+1)	3(+1)				
7장	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)			
8장	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)		
9장	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	2(+1)	
10장	2	2	2	2	2	2	2	2	2	2

```

function 문제출제(문제수, 예상평균점수, 단원수)
    난이도별출제갯수(5,2) = "난이도", '개수'를 초기화
    단원별출제문제수(단원수) = 0          ①
    단원별출제갯수 = 문제수 / 단원수      ②
    총출제문제수 = 0;
    for i = 0 to 4  ' 5단계의 난이도
        출제난이도 = 난이도별출제갯수(i,0)
        레코드셋 = select from 과목테이블 where 출제난이도
        출제문제수 = 난이도별출제갯수(i,1)
        count = 0
        do while(count < 출제문제수)
            난수(레코드번호) = int((total_count-1 + 1) * rnd() + 1 )
            if (단원별출제문제수(레코드 단원) < 단원별출제갯수)
                or ( 총출제문제수 >= 단원수/단원별출제갯수)
                exit()                                ③
            end if
        loop
    end for
end function

```

그림 2 난이도/영역별 문제출제 알고리즘

Fig. 2. question selection algorithm by Degrees of difficulty/domains

이에 표 4와 표 5를 기준으로 예상평균점수에 따른 난이도의 문제들을 출제함에 있어 어느 하나의 단원에 치우치지 않도록 전체 학습 범위에서 골고루 출제하는 과목별 자동 문제 출제 알고리즘은 그림 2와 같다. 알고리즘에서 ①은 단원수만큼 배열을 잡아서, 각 단원의 출제 개수를 체크한다. ②는 각 단원의 출제 개수를 계산한다. ③은 5개의 난이도에 대해 루프를 돌면서 문제를 출제할 때, 랜덤하게 선택된 레코드가 속해있는 단원을 체크해서 출제 가능 단원이면 루프를 빠져나가 시험 데이터베이스에 출제되고, 출제 불가능한 경우에는 계속해서 랜덤하게 문제를 추출한다.

### 3.2.3 재학습에 따른 재평가

특정 학습 영역에 대한 재학습의 결과를 평가하는 재평가에서는 해당 문제와 관련이 있는 문제만을 출제해야 하는 특징이 있다. 같은 학습 내용이라 해도 여러 단원에서 설명되어지는 경우와 현재의 학습 내용과 반드시 연계되어 설명되어져야 하는 선행 학습 영역의 경우에는 같은 개념이나 영역의 문제가 여러 단원에 존재한다. 그러므로 특정 학습 영역과 관련된 문제는 특정 단원에 한정된다고 볼 수 없으므로, 재평가를 위해서는 문제의 난이도와 각 문제의 키워드를 고려해서 전체 단원에서 특정 학습 내용과 관련된 문제만을 출제하는 난이도/키워드별 출제 알고리즘을 적용한다.

난이도에 대한 출제 문제수 비율은 표 3을 바탕으로 총 10문제를 출제 한다. 표 6은 총 10문제 출제에 대해 각 예상평균점수에 따른 난이도별 출제 문제수 비율을 계산한 결과이다. 다음으로 문제의 키워드를 고려함에 있어서는 틀린 문제의 ‘키워드1’, ‘키워드2’, ‘키워드3’ 필드값을 기준으로 한다. 그림 3은 표 6과 문제의 키워드를 기준으로 한 재평가를 위한 자동 문제 출제 알고리즘이다. 알고리즘에서 ①은 현재 출제하고자 하는 난이도와 키워드1, 키워드2, 키워드3에 맞는 문제들만으로 레코드셋을 구성한다. ②에서는 구성한 레코드셋의 개수가 현재 난이도의 출제 개수보다 적을 경우 랜덤하게 추출할 필요없이 모두 시험 데이터베이스에 출제하고, 다음 난이도 출제를 위해 루프를 돈다. ③에서는 레코드셋의 개수가 출제 가능 개수보다 많을 경우 랜덤하게 추출하여 가능한 개수만큼 시험 데이터베이스에 출제한다.

표 6. 재평가를 위한 난이도별 출제 문제수  
Table 6. The number of selection question by Degrees of difficulty for re-test

	예상평균 점수	총 시험문제 (10문제)				
		상	상중	중	중하	하
출제 비율 (출제 문제수)	90점	0	0	0.7(1)	2.4(2)	6.9(7)
	80점	0	0.2(0)	1.4(2)	3.4(3)	5.0(5)
	70점	0	0.7(1)	2.4(2)	4.2(4)	2.7(3)
	60점	0.2(0)	1.4(2)	3.4(3)	3.6(4)	1.4(1)
	50점	0.7(1)	2.4(2)	3.8(4)	2.4(2)	0.7(1)
	40점	1.4(2)	3.6(4)	3.4(3)	1.4(1)	0.2(0)
	30점	2.7(3)	4.2(4)	2.4(2)	0.7(1)	0
	20점	5.0(5)	3.4(3)	1.4(2)	0.2(0)	0
	10점	6.9(7)	2.4(2)	0.7(1)	0	0

```

function 문제출제(문제수, 예상평균점수, 키워드1, 키워드2, 키워드3)
    난이도별출제갯수(5,2) = "난이도별 출제문제수 비율" 테이블을
    참조하여 예상평균점수에 따른 '난이도' , '개수'를 초기화
    for i = 0 to 4   '5단계의 난이도
        출제 난이도 = 난이도별출제갯수(i,0)
        레코드셋 = select from 과목테이블 where ①
        출제난이도AND(키워드1OR키워드2OR키워드3)
        출제문제수 = 난이도별출제갯수(i,1)
        if(record_count <= 출제문제수) ②
            모든 문제를 사용자 시험 데이터베이스에 출제
            continue;
        end if
        count = 0
        do while(count < 출제문제수) ③
            난수(레코드번호) = int((total_count-1 + 1) * rnd() + 1)
            해당 레코드 문제를 사용자 시험 데이터베이스에 출제
            count = count + 1
        loop
    end for
end function

```

그림 3 난이도/키워드별 문제출제 알고리즘  
Fig. 3. question selection algorithm by Degrees of difficulty/keywords

## IV. 실험 결과 및 분석

이전의 연구[5]에서 구현하여 현재 일부 운영중인 웹 기반 학습 시스템에서 기존의 난이도만을 고려한 문제 출제 알고리즘과 본 논문에서 제시하는 평가 유형별 서로 다른 문제 출제 알고리즘을 각각 실험하여 그 결과를 비교 분석하였다.

먼저, 단원별 평가는 기존 알고리즘[4]과 같이 특정 단원에서 난이도만을 고려하기 때문에 비교 분석에서 제외하였다. 다음으로 과목별 평가의 출제 분포도를 분석하기 위하여 전체 9장으로 구성된 ‘운영체계’ 과목에 대하여, 기존의 알고리즘(알고리즘 A)과 본 논문의 알고리즘(알고리즘 B)을 각각 이용하여 예상 평균점수 50점

에 맞춘 난이도를 적용하여 문제를 출제하였다. 출제 횟수는 5번으로 한정하였고 매 출제 때마다 출제된 20개의 문제들에 대하여 어느 단원에서 출제되었는지 그리고 어떤 난이도가 출제되었는지에 대하여 비교하였다. 표 7과 표 8은 출제 문제에 대한 단원의 분포도 및 난이도의 분포도를 비교 분석한 결과이다.

표 7. 알고리즘 A의 출제현황  
Table 7. Selection state of algorithm A

	1장	2장	3장	4장	5장	6장	7장	8장	9장
1회	개수	0	3	4	7	3	0	0	2
	난이도		상1중1 중하1	하1 상중1 중2 중하2	상중2 중3 중하1	상중2 중3 중하1			중1 중하1 중1
2회	개수	2	7	5	0	1	2	0	3
	난이도	상1 중1	상중1 중4 중하2	상1 하1 상중1 중2		중하1	중하2		중1 상중2
3회	개수	6	0	0	3	0	4	1	2
	난이도	상1 중2 중하3			하1 상중1 중1		중3 중하1	중하1	상중1 중1 상중3 중1
4회	개수	3	1	1	2	1	3	7	0
	난이도	상1 중2	중1	상중1	하1 중1	상중1	상중1 중1 중하1	중하4 중1 상중2	
5회	개수	4	2	0	2	5	1	1	2
	난이도	중상2 중하1	중상1 중하1		상1 하1	중상1 중3 중하1	중1	중1 중하1 중1	중2 중상1

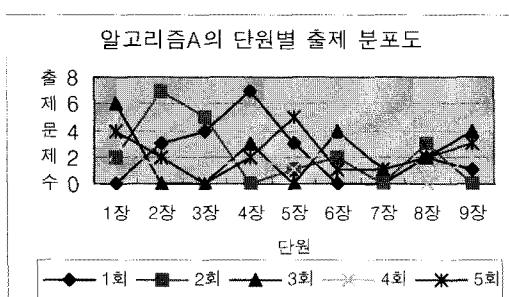


그림 4 알고리즘 A의 단원별 출제수 그래프  
 Fig. 4. The number of selection graph by units of algorithm A

표 7에서 알고리즘 A의 출제 문제의 난이도 분포는 예상 평균 점수에 맞추어 적절하게 출제된 반면, 출제 단원의 분포는 특정 단원에서 집중적으로 많이 출제되거나 한 문제도 출제되지 않는 단원도 나타남을 확인할 수 있다. 반면 표 8의 알고리즘 B 경우에는 난이도 분포가 예상 평균 점수에 맞추어져 있을뿐만 아니라, 어느 하나의

단원에 치우치지 않고 전 단원에서 골고루 출제되었음을 확인할 수 있다. 이는 특정 과목에 대한 학습 결과를 평가하는 시험에서는 과목의 전체 단원에서 골고루 문제를 출제하여야 하는 특징에 비추어 볼 때, 기준의 출제 알고리즘보다 학습자들에게 보다 알맞은 문제 출제 알고리즘임을 확인할 수 있다.

표 8. 알고리즘 B의 출제현황  
Table 8. Selection state of algorithm B

	1장	2장	3장	4장	5장	6장	7장	8장	9장
1회	개수	2	2	2	2	3	2	3	2
	난이도	상1 하1	상중1, 중하1	중2	상중1 중하1	상중2, 중1	중1 중하1	중하1 상중1 중1	중1 중하1
2회	개수	2	2	3	3	2	2	2	2
	난이도	중하2	상1 중1	상중2 중1	하1 중2	상중1 중1	상중1 중1	중2	중하2 상중1
3회	개수	3	2	2	3	2	2	2	2
	난이도	중3	상1 중1	하1 상중1	상중2 중1	중1 중하1	상중1 중1	상중1 중1	중1 중하1
4회	개수	2	2	2	2	2	3	2	2
	난이도	상중1 중하1	상중1 중1	중2	중1 중하1	상중1 중1	상1 상중1 중하1	중하1 중1	중1 하1 상중1
5회	개수	2	2	2	2	3	2	2	2
	난이도	중2	중상1 중1	중상1 중1	중상2	중1 중하1	중상1 중1	중1 상1	중1 하1 중1

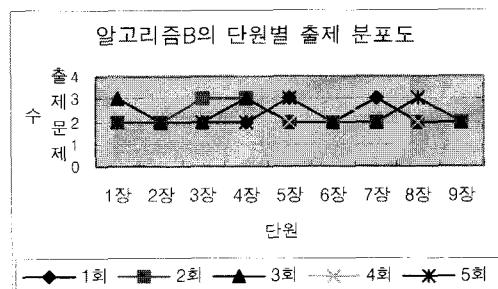


그림 5 알고리즘 B의 단원별 출제수 그래프  
Fig. 5. The number of selection graph by units of algorithm B

마지막으로 재평가를 위한 출제 분포도를 분석하기 위해 위의 과목별 모의시험 결과 5개의 틀린 문제에 대하여 재평가를 위한 문제를 출제하였다. 실험을 위하여 틀린 문제중에서 1장(문제a), 3장(문제b), 5장(문제c), 7장(문제d), 9장(문제e)에서 하나씩 선택하였다. 선택된 5개의 문제에 대하여 본 논문의 나이도/키워드별 알고리즘(알고리즘C)을 이용하여 예상 평균점수 50점에 맞춘 나이도를 적용. 재평가를 위한 문제를 출제하

여 어느 단원에서 출제되었는지, 어떤 난이도가 출제되었는지에 대해 분석하였다. 표 9는 알고리즘 C의 분석 결과이다.

표 9. 알고리즘 C의 출제현황  
Table 9. Selection state of algorithm C

단원문제	1	2	3	4	5	6	7	8	9
a 개수	10								
	난이도	상1, 상중2, 중4 중하2, 하1							
b 개수		3	7						
	난이도	중2 상중1	상1, 상중1, 중2 중하2, 하1						
c 개수		2	1	2	5				
	난이도	중1 중하1	중1 하1	상1, 상중2, 중2					
d 개수	2				2	6			
	난이도	중하1 하1			상중1 중1	상1, 상중1, 중3 중하1			
e 개수	3								7
	난이도	상1, 중1, 중하1				상중2, 중3, 중하1, 하1			

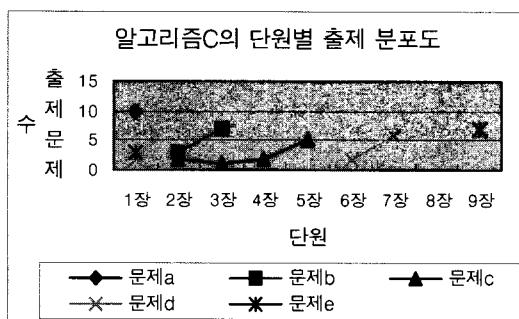


그림 6 알고리즘 C의 단원별 출제수 그래프  
Fig. 6. The number of selection graph by units of algorithm C

표 9에서, 문제a의 경우는 키워드가 ‘운영체제의 발전 과정’, ‘운영체제의 개념’으로 관련된 문제가 1장에 모두 한정되어 있고, 문제c의 경우에는 키워드가 ‘페이지 대치 알고리즘’, ‘가상메모리 관리’, ‘프로세스’로써 관련 문제가 2장부터 5장사이에서 출제되었음을 확인할 수 있다. 이는 재평가에서는 해당 문제와 관련 있는 내용의 문제가 출제되어야 하는 특징과 특정 내용과 관련 있는 문제들이 전체 단원에 분포되어 있을 수 있는 특징에 비추어볼 때, 재학습에 대한 평가를 원하는 학습자들에게 보다 알맞은 문제 출제 알고리즘임을 확인할 수 있다.

## V. 결론

이러닝 학습시스템에서 자기주도적 학습을 진행하는 학습자들이 학습 결과를 스스로 평가하는 모의시험의 효율성을 위해, 평가 종류에 따라 서로 다른 알고리즘을 적용함으로써 출제되는 문제들이 평가 종류에 최대한 적합할 수 있는 새로운 문제 출제 알고리즘을 제시하였다. 이는 평가의 종류가 갖는 특징을 비추어볼 때, 난이도만을 고려한 기존의 알고리즘에 비해 보다 효율적임을 실험을 통해 확인할 수 있었다.

## 참고문헌

- [1] 정용기, 최은만, “웹 기반 학습평가 자동화시스템의 설계 및 구현”, 한국정보처리학회논문지, 2002.
- [2] 하일규, 강병욱, “문항출제와 문항분석이 가능한 웹 기반 교육평가시스템의 설계및구현”, 한국정보처리학회논문지, 2002.
- [3] 임희숙, “웹기반 지능형 문제운행 시스템의 설계및 구현”, 전남대학교 석사학위논문, 1999.
- [4] 김경아, 최은만, “웹기반교육에서의 자동문제출제시스템”, 한국정보처리학회논문지, 2002.
- [5] 류희열, 김은정, “기사 시험을 위한 웹기반학습시스템의 설계”, 한국정보과학회 봄학술발표논문집, 2004.
- [6] 신민웅, “알기쉬운 통계학”, 생능출판사, 1995.
- [7] 김은정, 이상관, 김성곤, “이러닝 문제은행기반 출제 시스템을 위한 동적 난이도 조정정책”, 한국해양정보통신학회 논문지, 2008.

## 저자소개



김은정(Eun-Jung Kim)

1996년 경상대학교(공학석사)  
2001년 경상대학교(공학박사)  
1990~1993년 LG전자 멀티미디어  
연구소 연구원  
2001~2005년 부산외국어대학교 전자컴퓨터공학부  
초빙교수, 전임강사(강의전담)  
2006~2007년 부산대학교 시간강사  
2008~현재 부산가톨릭대학교 컴퓨터공학과전임강사  
<관심분야>정보검색, 원격교육, 웹기반 가상학습