

문제분석을 이용한 웹 기반 평가시스템

°허은주, 한현구

°한국의국어대학교 교육대학원 전자계산교육전공, 한국의국어 대학교 컴퓨터공학과

°zang1012@hufs.ac.kr, hghan@hufs.ac.kr

A Web-based Evaluation System with Query Analysis.

°Heo, Eun-Ju, Hyungoo Han

Hankuk University of Foreign Studies, the Graduate School of Education, Dept. of Computer Education
Dept. of Computer Science & Engineering, Hankuk University of Foreign Studies

요 약

정보사회와 컴퓨팅환경의 발전으로 언어, 학력, 인지도 등의 평가도구로서 컴퓨터기반의 평가환경이 보편화되었다. 특히 컴퓨터기반의 학력평가 환경을 위해서는 문항의 난이도와 변별도 등 문항의 특성을 정확히 분석하는 것이 필수적이다. 문항분석은 컴퓨터기반의 평가를 위한 양질의 문제은행 구성 및 문항과 수험자 능력의 정확한 추정과 체계적이고 과학적인 평가를 위한 전제조건이라 할 수 있다. 본 논문에서는 고전평가이론의 문제분석을 적용한 문제은행 시스템을 구현하였으며 수험결과를 분석하여 각 문제의 곤란도나 변별도, 문항분포도를 통하여 문제를 분석할 수 있도록 하였다. 또한 수험자가 각 문제를 푸는데 걸린 시간을 기록하여, 수험자의 문제에 이해도를 정확히 분석하고 수험자의 추측, 랜덤 선택 등으로 인한 정답을 맞힐 가능성과 한 문제를 읽고 이해하는 시간이 너무 오래걸린 이유에 대해서도 추정하였다. 문제분석 및 수험결과와 평가 및 분석으로 교사들은 문항의 양호도를 높일 수 있고 문제은행에 저장되어 있는 문항들을 수정하고 보완하여 양질의 문항을 출제할 수 있도록 하였다.

1. 서 론

컴퓨터의 발전과 더불어 평가도구로서 컴퓨터를 이용한 컴퓨터기반의 평가가 보편화되었다. 컴퓨터기반의 평가환경을 실현하기 위해서는 문제의 난이도와 변별도 등 문항의 특성을 정확히 분석하는 것이 필수적이며, 특히 문항분석은 컴퓨터기반 평가를 위한 양질의 문제은행 구성 및 문항과 수험자 능력의 정확한 추정과 체계적이고 과학적인 평가를 위한 전제조건이라 할 수 있다.

표준화된 학력검사나 공신력 있는 연구기관에서 실시하는 학력검사는 문항분석을 해서 만족할 만한 수준의 양호도가 확보된 문항만을 사용한다. 그러나 학교 현장에서 사용되는 대부분의 문제들은 문항분석을 하지 않고 있는 실정이다. 문항분석에 근거하지 않은 검사에 대한 평가는 주관적일 수밖에 없으며 검사의 질을 향상시킬 수 없다.

본 논문에서는 고전검사이론의 문항분석을 적용한 문제은행 시스템을 구현하였다. 학생이 각 문항에 반응한 시간을 기록하여 응답자가 문항에 대한 정답을 모르면서 추측이나 시험 보는 요령을 이용해서 답을 맞힐 가능성과 문제에 사용된 개념이나 용어에 대한 이해력 부족으로 인해 소모된 시간을 추정하였다. 이러한 추정치를 평가결과 분석에 적용하여 문항제작과 수정에 타당성을 높이고 문제관리가 효율적으로 이루어지도록 한다.

2. 관련연구

2.1 문항분석의 필요성

평가도구의 질은 그 도구를 구성하고 있는 문항의 특성에 의존하기 때문에 효과적인 평가도구를 만들기 위해서는 평가문항을 분석하여 검토하는 것이 필요하다. 평가문항이

원래 의도한 평가도구의 목적을 제대로 수행할 수 있도록 만들어졌는지를 다양한 측면에서 확인하는 작업을 문항분석 또는 문항의 양호도 분석이라고 한다.

문항분석이 필요한 이유는 첫째, 기본적으로 문항을 개선하기 위해서 필요하며 둘째, 문항분석 결과를 교수-학습 및 평가과정에 환류시킴으로써 교수-학습과정을 향상시키기 위하여 필요하며 셋째, 문항분석은 교사와 학생들이 수험결과에 대한 토론 과정의 유용도와 효율성 제고에 필요하며 넷째, 문항분석을 통한 교사의 전문성향상에 필요하며 다섯째, 문항은행의 구축에 필요하다. [1][2]

2.2 문항분석 방법

2.2.1 문항곤란도

문항곤란도(item difficulty)란 한 문항의 어려운 정도를 뜻하며 문항난이도라고도 불린다. 문항 곤란도 지수(item difficulty index)는 <표 1>과 같이 각 문항에 대하여 반응한 사람수에 대하여 정답으로 반응한 사람 수의 백분율이나 비율로 표시된다.[6]

<표 1> 문항곤란도

$P = 100 \times \frac{R}{N}$	P = 문항곤란도 지수
	N = 전체 사례수
	R = 정답 학생수

2.2.2 문항변별도

문항변별도(item discrimination)란 어떤 평가의 개개 문항이 그 평가에서 득점이 낮은 학생과 높은 학생을 식별 또

는 구별해 줄 수 있는 변별력(discrimination power)으로서 <표 2>는 문항변별도의 한 방법인 육간상관표에 의한 문항 변별도 산출방법이다.[2]

<표 2> 문항변별도

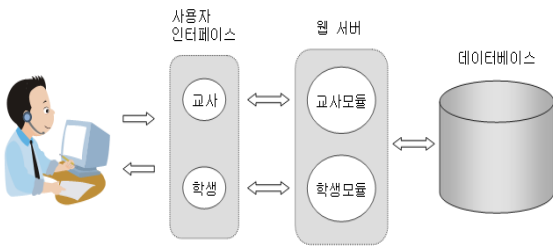
$D.I. = \frac{HR-LR}{\frac{N}{3}}$	D.I. = 문항변별도 지수
	HR = 상위집단의 정답자 수 LR = 하위집단의 정답자 수 N = 사례수

2.2.3 문항반응분포

문항반응분포는 문항 속에 포함되어 있는 답지 하나 하나에 학생들이 어떻게 반응했는지를 분석하는 것이다. 이를 통해 문항의 답지가 의도했던 기능이나 역할, 즉 각 답지의 오답이 얼마나 오답으로서의 매력에 있으며, 정답은 얼마나 정답 구실을 했는 가 등을 알아보는데 필요한 것이다.[5]

3. 시스템의 설계

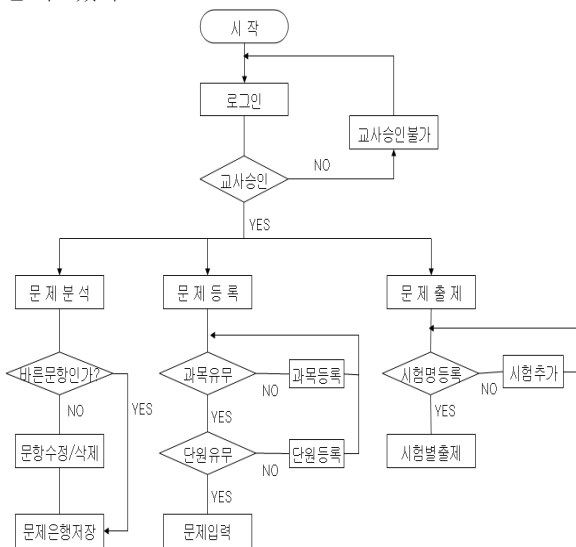
본 논문에서 구현한 시스템 구조는 <그림 1>과 같다.



<그림 1> 평가시스템 구조

3.1 교사모듈

교사 모듈은 <그림 2> 교사모듈 전체흐름도로 상세히 설명될 수 있다.



<그림 2> 교사모듈 전체 흐름도

3.1.1 문제 분석

문제 분석 기능은 본 시스템에서 가장 큰 부분을 차지하는 것으로 분석의 종류를 크게 네 부분으로 나누어 설계하였다.

첫째, 곤란도(난이도)를 구하는 기능으로 응시한 전체 학생 수에 해당문제의 정답자 수를 나누어 계산하였다. <표 3>과 같이 난이도의 수치가 높을수록 쉬운 문항으로 판단할 수 있다. 문제 등록에 입력되었던 교사의 예상난이도와 학생이 응시한 시험결과를 이용해 분석한 결과를 비교하여 교사가 문제를 수정하거나 삭제할 수 있도록 하였다.[4]

<표 3> Cangelosi 의 문항난이도에 의한 문항평가

난이도	문항평가
25 이하	어려운 문항
25 ~ 75	적절한 문항
75 이상	쉬운 문항

둘째, 학생이 응시했던 시험의 결과를 데이터베이스에서 검색하여 전체 성적에서 상위 27% 집단과 하위 27% 집단으로 분류한 다음 각 문제의 정답자 비율이 어느 쪽이 더 많은지를 D.I.지수를 통해 구하였다.

결과에서 상위집단이 하위집단 보다 많은 정답자가 나왔다면 그 문제는 학생들의 상·하 능력을 변별해줄 수 있는 문제라고 할 수 있다. 반면에 상위집단보다 하위집단에 더 많은 정답자가 나온 문제라면 그 문제는 변별력이 부족하다고 판단할 수 있다.[4]

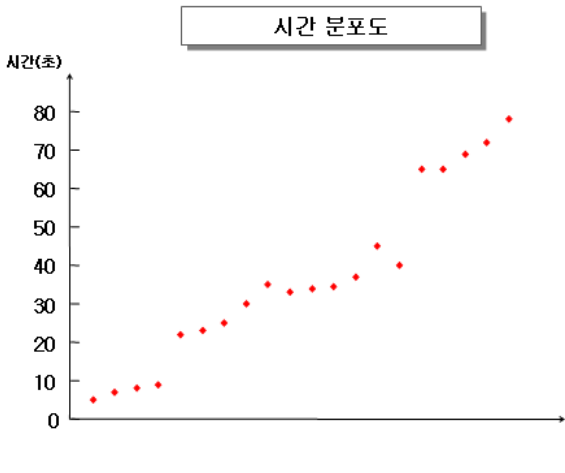
셋째, 문제의 오답지 분포도(오답지 매력도)는 정답을 제외한 오답에 분포된 학생 수를 그래프 형태로 표시할 수 있도록 하였다. 매력적인 오답이 되기 위해서는 오답에 응답한 학생들의 분포가 비슷해야 한다. 오답지의 분포가 비슷하다는 것은 문제의 난이도가 높아질 수 있다고 볼 수 있다.[3]

넷째, 한 문제를 푸는데 걸리는 시간에 대한 분포도이다. 컴퓨터화면에서 한 문제씩 풀 수 있도록 하여 문제가 로딩된 시점에서 다음 문제를 풀기 위한 버튼을 누를 때까지의 시간을 저장한다. 이전에 풀었던 문제를 수정하기 위해 돌아가는 경우 저장되어 있던 시간에 다시 푸는데 소요된 시간이 누적되도록 하였다.

<그림 3>은 95학년도 대학수학능력시험 수리탐구영역II (예·체능계) 에 출제되었던 문제에 대한 분석 결과이며 다음과 같이 평가할 수 있고 <표 4>와 같이 해석할 수 있다.

첫째, 10초 이내에 풀어서 정답을 한 학생은 문제를 읽는데 걸리는 최소 시간보다 적기 때문에 성실하게 문제를 풀었다고 볼 수 없으며 이 중 정답자가 있었다면 추측에 의해 맞은 경우로 볼 수 있다.

둘째, 제한시간인 60초를 초과하여 응답한 학생은 문제의 난이도가 높거나 문제에 포함된 개념이나 관련용어가 어려웠다고 볼 수 있다. 지문이 너무 길어 학생에게 지루함을 줄 수 있고 불필요한 단어로 인해 혼란을 가져올 수 있다.



<그림 3> 예상시간분포
<표 4> 반응시간 분포도에 대한 해석

시간(초)	정답자	오답자
10 이하	추측에 의한 정답. 난이도 낮음.	성실하지 않은 응답자. 전체 시간이 부족함.
20 - 50	정상 분포	학습능력 부족
60 이상	난이도 높음	개념이나 용어에 대한 이해력 부족

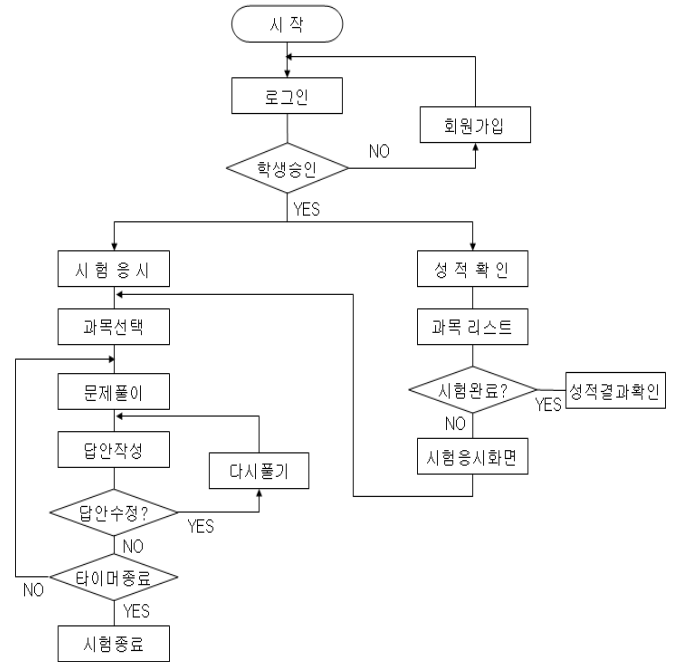
3.1.2 반응시간과 기존 분석 방법과의 관계

문제의 난이도는 단순히 전체학생 수에 정답자 수를 나눈 값을 비율로 나타내고 있다. 단순한 공식에 의한 결과로 문제를 분석하는 것은 한계가 있다고 할 수 있다. 교사는 문제를 출제할 때 학생이 문제를 푸는데 걸리는 평균적인 시간이 40초 이상 소요될 것이라고 판단했으나 반응시간의 결과는 대부분의 학생들이 30초 이내에 응답을 한 결과가 있다고 가정할 때, 문제의 답항이 제 기능을 못하고 있거나, 학생의 수준을 제대로 파악하지 못한 경우로 볼 수 있다. 이처럼 문제에 대한 반응시간을 측정하게 되면 좀 더 구체적인 결과를 파악할 수 있다.

난이도분석의 경우 전체 학생 수에서 답을 하지 않은 학생 수를 제외하고 계산하는 방식이 있는데 시간이 부족해서 답을 선택하지 않은 문제는 오답자로 처리하지 말고 학생 수에서 제외하자는 것이다. 이때 마지막에 연속해서 못 푼 문제는 시간부족으로 인한 미달항으로 판단할 수 있지만 풀다 못 푼 문제(제외항), 풀어보지도 않고 넘긴 문제(월번항)는 육안으로 구별하기가 힘들다. 그러나 웹기반 평가에서는 풀다 어렵거나, 시간이 부족할 때 다음문제로 넘어갈 수 있는 기능이 있기 때문에 여기에 반응시간 결과를 두 경우에 대입하면 제외항과 월번항에 대해서 구분할 수가 있으며 이를 본 논문에서 활용하였다.

3.2 학생모듈

학생 모듈은 <그림 4> 학생모듈 전체흐름도로 상세히 설명될 수 있다.



<그림 4> 학생모듈 전체 흐름도

3.2.1 시험 응시

학생은 회원 가입한 아이디로 로그인하여 시험에 응시할 수 있다. 시험명과 과목명을 선택하여 시험에 응시하고 학생이 입력한 정답은 답지에 표시되며 데이터베이스에 저장된다. 모든 문제를 다 풀었거나 시험 제한시간이 종료되면 답지가 제출되도록 하였다.

3.2.2 성적 확인

학생은 과목리스트에서 자신이 응시한 과목명을 선택하여 성적 결과를 확인할 수 있도록 하였다. 만약 응시하지 않은 과목이 있다면 해당 과목명을 선택하여 시험보기 화면으로 돌아갈 수 있도록 구성하였다. 과목에 대한 총점과 자신이 응답한 답항과 문제의 정답을 색상으로 구분하여 표시하고 해설도 함께 보일 수 있도록 하여 문제에 대한 피드백을 바로 받을 수 있도록 설계하였다.

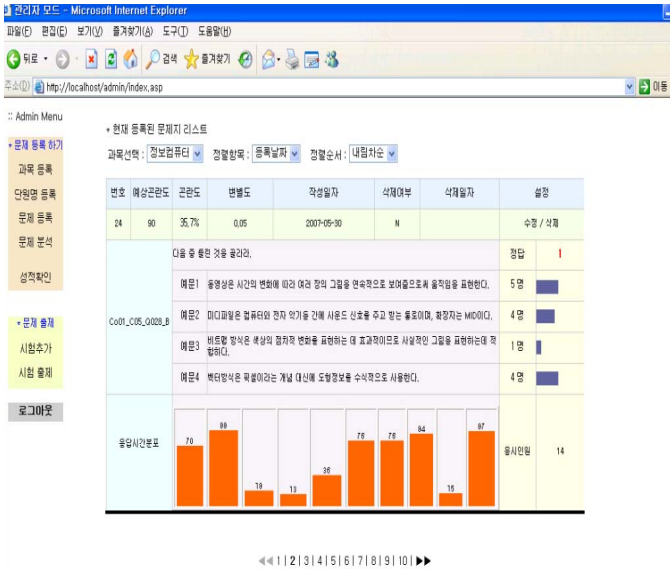
3.3 데이터베이스

데이터베이스를 구현하기 위하여 DBMS로 MS-SQL Server를 사용했으며, 데이터베이스의 명칭을 SAT로 하고 데이터베이스를 생성하였다. 문제를 등록하기 위해 Lesson 테이블, Category 테이블, Question 테이블, FilePath 테이블로 구성하였고 문제를 출제하기 위해 Test 테이블과 Question_Setting 테이블을 두었다. 또한 학생의 정보를 저장하기 위해 Student 테이블과 시험에 출제되었던 문제별로 학생의 아이디, 답지, 시험날짜와 문제번호, 문제당 시간 등의 정보를 기록하는 Personallog 테이블로 구성하여 문제별 분석에 이용될 수 있도록 설계하였다.

4. 시스템의 구현

다음 그림은 본 시스템에서 가장 핵심인 문제 분석화면

이다. 평가를 마친 문항에 대한 데이터를 계산하여 곤란도, 변별도를 계산하고 <그림 5>와 같이 오답지에 대한 분포와 한 문항을 푸는데 소요된 시간분포를 그래프형태로 보여주도록 하였다.



<그림 5> 문제 분석 화면

5. 결론

학교에서 주로 이루어지는 평가는 학생들의 학업성취에 대한 평가이다. 이를 위해서는 해당 표준화 학력검사를 이용할 수도 있지만, 대개는 교사가 만든 필답고사를 이용하여 평가하는 것이 일반적이라 할 수 있다. 하지만 새로운 문항의 출제에 걸리는 시간과 노력, 채점에 걸리는 시간들을 고려해볼 때 지필고사보다는 웹을 통한 평가가 효율적이라고 볼 수 있다.

본 논문에서 구현한 시스템은 다음과 같은 장점이 있다고 할 수 있다. 첫째, 교사는 본 시스템을 통해 분석결과를 쉽게 확인할 수 있고 다음 문항제작에 참고할 수 있도록 하였다. 분석방법을 단순하게 적용하기 보다는 반응시간분포도에서 표준시간대에서 벗어난 대상을 기존의 난이도나 변별도, 오답지 매력도와와의 관계를 파악하여 보다 정확한 분석결과를 유추할 수 있다. 둘째, 학생은 본 시스템을 통해 시험을 응시하면 바로 시험 결과를 확인할 수 있고 틀린 문항에 대해 즉각 피드백을 받을 수 있어 보다 빠르게 학습 향상이 되도록 하였다.

문제해결 시간 측정 외에 좀 더 다양한 방법을 개발하여 양질의 문제를 구성할 수 있는 시스템이 될 수 있도록 본 시스템을 향상시키는 작업이 앞으로 필요하다 할 수 있겠다.

[참고문헌]

- [1] 강승호 외, 현대교육평가의 이론과 실제, 양서원, 2004
- [2] 정송진, 학교학습의 극대화를 위한 교육평가의 이해, 양서원, 2004
- [3] 성태제, 문항제작 및 분석의 이론과 실제, 양서원,

2004

- [4] 송인섭 외, 교육과정 및 교육평가, 양서원, 2001
- [5] 박도순, 교육평가 「이해와 적용」, 교육과학사, 2007
- [6] 이수미, 웹 기반에서의 문항분석을 통한 문제은행 시스템의 설계 및 구현, 홍익대학교 교육대학원, 석사학위 청구 논문, 2003
- [7] <http://www.kice.re.kr/kice>