

문항 특성을 고려한 문제 추출 컴포넌트 설계 및 구현

정 화 영

요 약

대부분의 웹 기반 학습은 학습자에게 획일적이고 일관된 학습자료를 제공해주고 있다. 이에 문항분석방법을 적용하거나 학습패턴을 이용하여 웹 기반 학습 시스템에 도입하려는 연구가 진행되었다. 그러나, 이러한 방법을 적용하기 위해서는 복잡한 알고리즘이나 파라미터 설정등이 필요하다.

따라서, 본 연구에서는 문항출제특성 중 문항에 대한 학습자의 오답율과 출제 빈도수를 고려하여 문제를 추출하는 시스템을 설계 및 구현하였다. 또한, 문제 추출에 관한 비즈니스 로직을 EJB로 구현함으로써, 효율적인 시스템 개발이 가능하였으며 유지보수 및 재사용성을 높였다.

A Design and Implementation of the Question Selection Component considering Item Attribute

Hwa-Young Jeong

ABSTRACT

Most web-based learning is furnishing single side and consistent training resource to learner. Research was gone to apply item analysis method or to introduce web-based learning system using studying pattern. But, we need complicated algorithm or parameter setting etc, for apply these method. Therefore, in this research, we design and implement an item selection system in consideration of learner's incorrectness rate and problem frequency selection rate about question of item selection attribute. Also, as that embody business logic about item selection by EJB, efficient system development is available and we improved maintenance and reusability.

1. 서 론

웹 기반 교육을 위한 콘텐츠는 수많은 잠재 고객을 가지고 있으며, 웹 기반 강의는 언제 어디서나 많은 수의 학습자에게 각자의 능력에 맞게 학습할 수 있다[1]. 즉, 웹은 교사와 학생간, 학생

들간 또는 학생과 사회 각 분야의 여러 사람들간에 활발한 정보교환이 이루어지도록 하여 학습공동체를 형성한다[2]. 인터넷을 이용한 문제 해결 학습은 학습자들간의 상호작용을 증가시키면서, 그들이 보다 많은 생각을 공유하면서 문제를 해결해 나가도록 하였다[3]. 그러나, 현재 사용되고 있는 웹 기반 코스웨어는 개개인의 학습능력에 따라 학습내용과 학습방법을 다르게 제시하는 개

별화 학습이 부족하다[4]. 또한, 학습내용을 가지고 있는 문제은행시스템에서 기존의 문제출제 방식은 대부분 고정출제나 무작위출제방식으로 보다 효율적인 학습내용을 제공하지 못하고 있다. 따라서 웹 기반 시스템 개발과정에 있어서도 정적인 시스템 개발에서 동적인 시스템 개발 방법이 요구되고 있다[5]. 기존의 연구에서 김경아[11]는 문항난이도를 고려한 문제출제방식을 제안하였다. 그러나 초기의 문항난이도 설정이 임의적이며, 대량의 문제은행 시스템에서 문항난이도만을 고려하는 것은 출제 문제 선택에 있어서 보다 효율적인 방법이 되지 못한다.

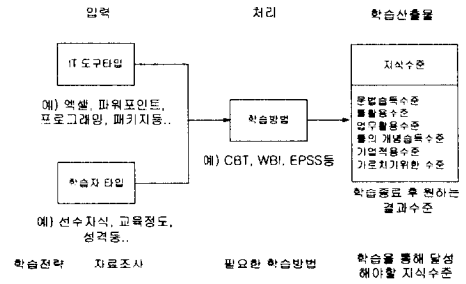
또한, 소프트웨어 개발부에서도 점차 복잡하고 다양해지는 소프트웨어의 품질향상과 시기적절한 소프트웨어 생산 및 요구사항 변화에 대한 효과적 대응을 위하여 소프트웨어를 재사용 가능한 컴포넌트 형태로 개발하는 기술이 빠르게 발전하고 있으며[6], 웹 기반 교육 시스템 분야도 이에 대한 적절한 대응이 요구된다.

본 연구에서는 학습자의 학습 효과를 높이기 위하여 웹상의 학습 출제문항 특성을 고려한 문제추출 시스템을 설계 및 구현하였다. 또한, 소프트웨어 개발 방법의 새로운 대안인 컴포넌트 기반으로 구현하였으며, 구현 모델은 서버측 컴포넌트 모델인 EJB를 이용하였다. 적용된 출제 문항특성은 학습자들의 오답율과 출제 빈도수를 고려하였으며, 웹 코스웨어에 추출 컴포넌트를 추가함으로써 본 시스템이 적용 가능함을 보였다.

2. 관련연구

2.1. 학습 프레임워크

학습자의 학습 및 훈련에 대한 전략적 모델로서 Sein, Bostrom and Olfman은 정보기술의 형태와 학습자의 타입에 따라 적절한 학습방법을 선정해야 하고 각각에 맞는 지식수준의 교육적 효과가 나타나야 한다는 모델을 <그림 1>과 같이 제시했다. 이는 학습자에게 적절한 학습방법을 찾아가는데 활용될 수 있다[7].



<그림 1> 학습전략 프레임워크

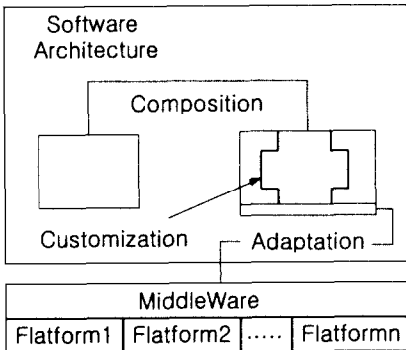
2.2. 문제 출제 시스템

웹 기반 문제 및 평가 시스템 영역은 사용자 인터넷을 통하여 시험을 보고 그 결과를 자동적으로 관리해주는 기능을 갖춘 시스템을 의미한다 [8]. 인터넷을 이용한 시험의 경우 정형화된 문제를 사용한다면 응시자 사이에 보안이 유지될 수 없다. 따라서 기존의 연구에서는 방대한 양의 문제 데이터베이스에서 조건에 맞는 문제를 선택하거나 문제의 패턴은 일정하되 그 패턴 안에서 사용되는 파라미터나 순서쌍을 바꾸는 출제 방법 등이 고려되었다. 그러나 문제 데이터베이스에서 조건에 맞는 문제출제 방법에서는 학습문제 추출을 위하여 출제 담당자의 전문가적인 조건입력이 필요하다. 또한, 파라미터나 순서쌍을 이용한 출제 패턴방식은 문제 추출 패턴에 따른 각 파라미터 설정등 추가적인 작업이 필요하며, 설정값 관리 및 변경 등에 따른 시스템 유지보수가 필요하다. 또한, 김경아[11]의 연구에서는 문항난이도를 고려한 자동문제 출제시스템이 제시되었다. 즉, 출제자가 난이도(상, 상중, 중, 중하, 하) 및 정답률(%)을 임의로 부여하고 학습과정에서 학습자의 점수에 따라 정답률 및 난이도를 재조정하면서 문제를 출제하도록 하였다. 그러나 위 방식은 출제자 및 학습자 측면에서만 고려되었다. 즉, 문제은행측면에서는 다양하고 대량의 학습내용을 가지고 있으며, 문제출제는 그 중에서 일부를 추출하여 학습자에게 제공된다. 따라서 각 문항들의 출제빈도수가 다르며 어떤 문제가 출제되느냐에 따라 학습효과도 달라진다. 학습자에 대한 다양하고 적절한 학습내용을 제공하기 위해서는 각 문제별 학습결과만이 아니라 문제출제 빈도수도

같이 고려되어야 한다.

2.3. 컴포넌트 기반 개발 방법

컴포넌트 기반 개발은 기존의 소프트웨어를 통째로 개발하던 기존의 방식과 달리, 부품 역할을 하는 컴포넌트를 기능별로 개발하고 이중 필요한 것만을 선택하여 조립하는 소프트웨어 개발 방법이다[9]. 이는, 크게 3가지 관점으로 나타낼 수 있다. 첫째, 이미 만들어진 컴포넌트를 이용하여 새로운 소프트웨어를 구성하는 통합자 관점. 둘째, 재사용될 수 있는 컴포넌트를 만드는 제공자 관점. 셋째, 제공자가 만들어 놓은 컴포넌트를 통합자가 효과적으로 이용할 수 있도록 도와주는 기능층 하는 중간자 관점 등이다. 컴포넌트 기반 개발방법(CBD:Component Based Development)에서 요구되는 주요활동은 <그림 2>와 같다[10].



<그림 2>CBD에서의 주요활동

3. 문항출제 특성을 고려한 문제추출 컴포넌트 설계

본 연구에서 적용한 문항출제 특성을 고려한 문제 출제방식은 기존의 학습효과만을 고려했던 방과는 다르게 문제출제 횟수를 같이 고려함으로써 문제에 대한 추출항목을 보다 더 정확히 반영하고자 하였다. 또한, 현재의 소프트웨어 개발형태인 컴포넌트 기반으로 구현함으로써 중요 알고리즘을 독립 기능단위 모듈로 분리시켜 시스템 개발의 효율성, 재사용성 및 유지보수성을 높였다. 학습자들의 오답율은 기존의 문항분석 방법

에서 문항난이도와 같은 방식으로 산출되며, 문제 출제 빈도수는 문제 데이터베이스에서 추출된 문제들의 출제율을 나타낸다. 즉, 본 연구에서는 문제 데이터베이스에 예제 50문제를 입력하였고 문제 출제 시에는 학습자들의 오답율과 출제 빈도수에 따라 10문제를 추출하여 학습을 진행하도록 하였다. 이를 위하여 데이터베이스는 Informix사의 Cloudscape를 이용하였으며, 문제 추출 기능모듈을 EJB로 구현하였고 화면구성 및 제어는 JSP와 나모5.0을 이용하였다.

3.1. 문제 데이터베이스 및 문항 추출 알고리즘

본 연구의 문제 데이터베이스는 <표 1>과 같다. 출제횟수는 전체 문제 중에서 해당 항목번호의 문제가 출제된 횟수를 나타내며, 전체 출제횟수는 항목번호의 문제와 상관없이 전체 응시횟수를 나타낸다. 즉, 예제 50문제 중에서 전체출제횟수가 30이며 1번 항목번호의 문제에 대한 출제횟수가 20이면, 학습자들이 본 시스템을 30번 이용하여 문제를 제공받았음을 뜻하고 그 중에서 1번 항목의 문제가 20번 추출되어 학습자에게 제공되었음을 나타낸다. 오답수는 출제된 항목번호에 대한 문제를 틀린 학습자들의 수이며 본 문제 데이터베이스를 이용한 전체 이용자의 수, 즉 전체 출제횟수에서 오답자들의 수를 나타낸다. 오답율은 전체 학습자 대비 오답자의 비율을 나타낸다.

$$\text{오답율} = \frac{\text{오답수}}{\text{전체출제횟수}} * 100$$

전체 출제횟수는 학습자들이 본 문제 데이터베이스를 이용한 횟수로서 전체 학습자들의 이용횟수와 일치한다. 출제율은 전체 출제횟수 대비 해당 항목번호 문제의 출제횟수 비율이며 출제 빈도수를 나타낸다.

$$\text{출제율} = \frac{\text{출제횟수}}{\text{전체출제횟수}} * 100$$

<표 1> 문제 데이터베이스 내용

Item	타입	비고
Num	숫자	항목번호
Contents	문자	문제내용
Exam 1	문자	예시 1번 내용
Exam 2	문자	예시 2번 내용
Exam 3	문자	예시 3번 내용
Exam 4	문자	예시 4번 내용
Correct	숫자	정답
Outcount	숫자	출제횟수
Incr_cnt	숫자	오답수
Incr_yld	숫자	오답율
Totoutcnt	숫자	전체출제횟수
Out_yld	숫자	출제율
Decision	문자	출제유무

출제유무는 산출된 오답율과 출제율을 기준으로 <표 2>와 같이 나타낸다.

<표 2> 문제 출제유무 결정 테이블

번호	문제 추출 조건	출제유무
1	오답율 >= 30 AND 출제율 >= 50	출제
2	오답율 >= 30 AND 출제율 < 50	출제
3	오답율 < 30 AND 출제율 >= 50	미출제
4	오답율 < 30 AND 출제율 < 50	출제

문제추출 조건에서는 오답율 추출조건 30%와 출제율 추출조건 50%로 설정하였다.

<표 3> 문항 난이도 및 오답율 설정 테이블

난이도	오답율(%)	비고
상	90 ~ 100	어려운 문제
상중	70 ~ 89	조금 어려운 문제
중	50 ~ 69	보통
중하	30 ~ 49	쉬운 문제
하	30 이하	아주 쉬운 문제

김경아[11]의 연구에서 난이도에 따른 정답율

설정기준을, 본 연구에서는 같은 방식으로 오답율을 기준으로 난이도에 따라 설정하면 <표 3>과 같다. 따라서 <표 2>의 오답율 추출조건에서 30%를 기준으로 한 것은, 오답율 30%이하의 아주 쉬운 문제는 학습효과를 기대하기 어려우므로 이를 제외한 나머지 문제가 오답율에 관한 출제 대상이 되도록 하였다. 같은 방식으로 출제율 설정기준을 나타내면 <표 4>와 같다.

<표 4> 문항 출제율 설정 테이블

출제빈도	출제율(%)	비고
상	90 ~ 100	빈번히 출제됨
상중	70 ~ 89	자주 출제됨
중	50 ~ 69	보통
중하	30 ~ 49	간혹 출제됨
하	30 이하	거의 출제 안됨

따라서 <표 2>의 출제율 추출조건에서 50%를 기준으로 한 것은, 문항출제가 간혹 또는 거의 출제되지 않는 출제빈도 중하·하와 보통이상으로 출제되는 중 이상의 기점이 50%이므로 출제율을 50%이하로 설정함으로써 보통이하로 자주 출제되지 않는 문제를 대상으로 하였다. <표 2>에서 1번의 경우, 오답율이 30%를 넘고 출제율도 50%를 넘음으로서 이는 자주 출제되면서도 학습자들이 많이 틀리는 문제항목으로서 출제유무에서는 출제로 판정한다. 3번의 경우는 자주 출제되지만 오답율이 적음을 나타내므로 미출제로 결정된다. 또한, 시스템의 초기 적용시 오답율과 출제율이 없거나 출제유무가 없는 문항에 대해서는 문제추출 알고리즘을 적용하지 않고 무작위로 문제를 출제하여 학습토록 한다. 만일, 출제유무에서 출제의 항목이 추출하여야할 문제수인 10문항 이하인 경우에도 출제항목을 우선으로 문제를 추출하되 나머지는 미출제 항목에서 무작위로 출제하였다.

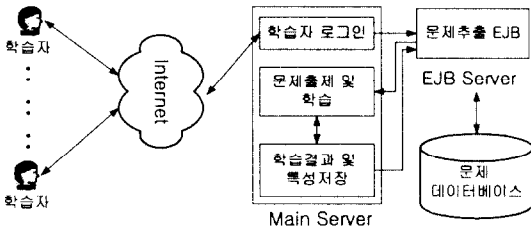
<표 5> 학습자의 학습정보

Item	타입	비고
Num	숫자	항목번호
Answer	숫자	답안번호

또한, 학습자의 학습결과를 나타내기 위한 임시 저장장소는 <표 5>와 같이 나타낸다.

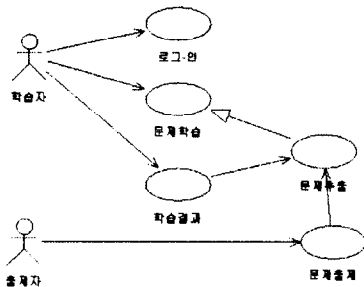
3.2. 문제추출 컴포넌트 설계

본 시스템은 메인 서버와 EJB 컴포넌트 서버를 갖는 분산서버로 구성하였으며 EJB부분은 무상태 세션빈으로 구현하였다. 메인 서버의 기능들은 HTML과 JSP로 구현하여 학습화면과 학습응답을 제어하며 문제추출에 관한 부분은 비즈니스 로직으로서 EJB서버에 두었고 문제 데이터베이스는 J2EE에 내재 되어있는 Informix사의 Cloudscape를 이용하였다. 이에 따라 구성은 <그림 3>과 같이 나타낸다.



<그림 3> 시스템 구성도

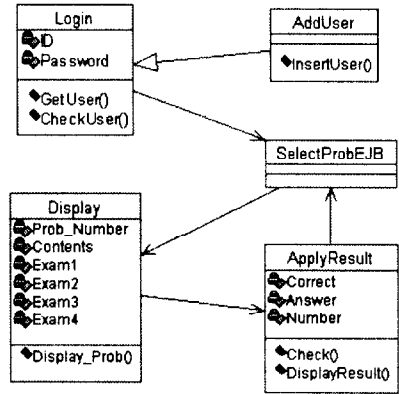
즉, 학습자 로그-인후 문제추출을 위하여 문제추출 EJB의 문제출제유무를 산출하는 문제선택 메소드를 호출하며 문항당 출제횟수를 가산하고, 출제될 문제가 나타나면 학습이 이루어진다.



<그림 4> Usecase Diagram

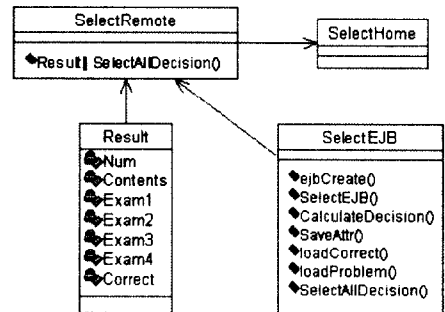
학습결과 및 특성저장 부분에서는 학습후에 정답확인을 하며, 오답수를 저장한다. 이에 대한 유즈케이스 다이어그램은 <그림 4>와 같다.

본 시스템의 각 클래스 관계는 <그림 5>에서 같이 구성하였다.



<그림 5> 전체과정에서의 Class Diagram

로그-인 이후에 SelectProbEJB부분에서는 문제추출을 위한 EJB컴포넌트를 등록 및 호출하며, Display부분에서는 추출된 문제에 따라 문항번호, 문제내용, 예시 항목등을 나타냄으로서 학습을 진행시킨다. 학습의 결과는 ApplyResult에서 나타내며 항목번호를 기준으로 정답과 답안을 검색하여 그 결과를 나타내며, EJB컴포넌트를 호출하여 다음 학습의 문제추출을 위해 오답수를 저장한다. 이에 대한, 문제추출의 비즈니스로직을 갖는 EJB컴포넌트 부분은 <그림 6>과 같다.

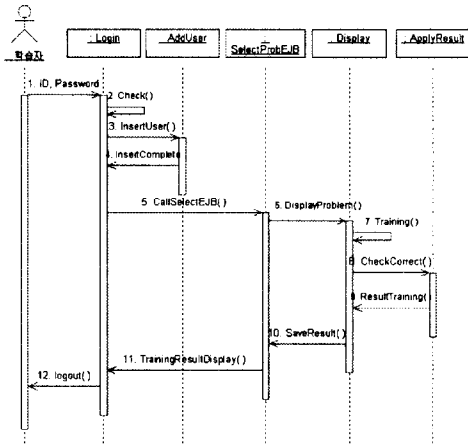


<그림 6> 문제추출 EJB 컴포넌트의 Class Diagram

무상태 세션빈으로 구현된 문제추출 EJB 컴포넌트는 EJB서버측면에서 클라이언트가 되는 메인 서버의 SelectProbEJB부분에서 SelectHome부

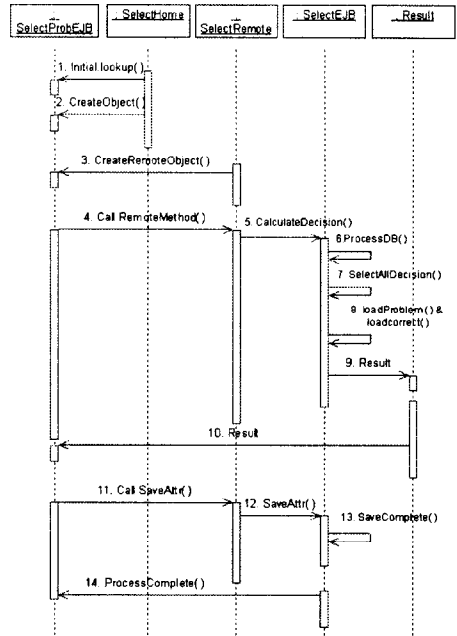
분을 호출함으로써 EJB를 사용할 수 있도록 등록 및 초기화하며, 이를 사용하는 리모트 인터페이스인 SelectRemote부분에서 실제 기능모듈을 실행하는 SelectEJB를 호출하여 그 결과를 Result클래스를 통하여 받는다.

이를 기반으로 전체과정에서 각 객체사이의 순서도는 <그림 7>과 같다.



<그림 7> 전체과정에서의 Sequence Diagram

로그-인 이후 SelectProbEJB를 통하여 문제를 출제하며 Display와 ApplyResult에서는 학습과정답을 확인하는 과정을 거치며, 다시 Display부분에서 SaveResult를 EJB 컴포넌트에 호출하여 오답수를 저장한다. 실제 문제 추출 기능과 오답수를 저장하는 EJB 컴포넌트 부분은 <그림 8>과 같다. 즉, EJB 컴포넌트를 사용하기 위하여 lookup을 통하여 접속 컴포넌트를 찾으며 CreateObject 메소드에서는 문제추출 EJB인 SelectEJB를 등록한다. 또한, 해당 EJB의 리모트 인터페이스에 대한 객체를 생성한 후 이를 통하여 문제추출 알고리즘을 갖는 SelectEJB의 CalculateDecision메소드를 호출하게 된다. 문제 데이터베이스에서 선택된 추출 문제들은 Result를 통하여 메인 서버에 나타내면 학습자가 학습을 시작할 수 있다. 학습과정에 따른 결과는 메인 서버 측에서 답안과 정답을 비교한 후 오답수를 EJB컴포넌트를 통하여 데이터베이스에 갱신한다.

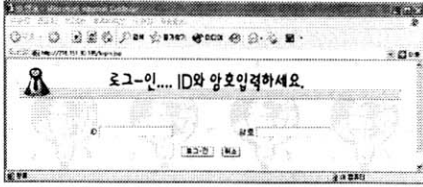


<그림 8> 문제추출 EJB 컴포넌트의 Sequence Diagram

4. 문항출제 특성을 고려한 문제추출 컴포넌트 구현

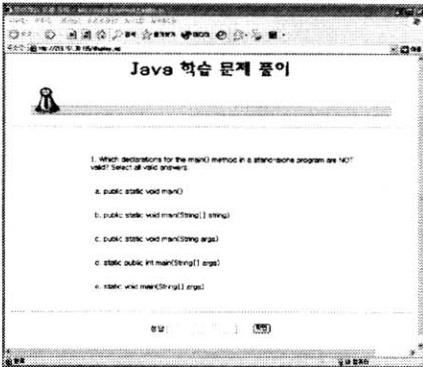
본 시스템의 구현 및 운용환경에서 운영체제는 Window XP를 이용하였고, 웹 서버언어로는 뷰를 위하여 HTML를, 학습자의 핸들링 및 비즈니스 로직을 위하여 JSP를 이용하였다. 또한, 비즈니스 로직부분 중에서 문제추출 부분은 J2EE를 이용한 EJB 컴포넌트로 구현하였으며, 데이터베이스는 Cloudscape를 이용하였다. 본 시스템의 문제학습을 위하여 SCJP자격증 취득을 위한 Java 문제풀이 예제 50문항을 출제하여 데이터베이스에 직접 입력하였으며 문제추출 알고리즘에 따라 10문항을 출제하도록 하였다.

이에, 본 시스템의 로그-인 화면은 <그림 9>와 같다.



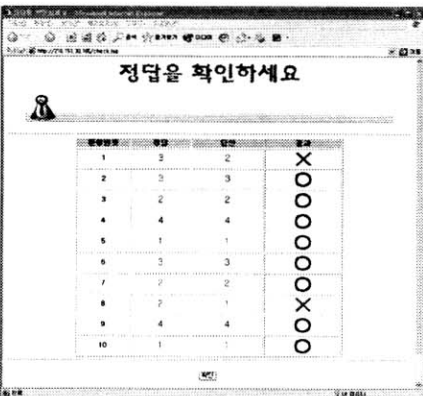
<그림 9> 로그인 화면

로그인후 <그림 10>과 같이 문제 데이터베이스의 50문항 중에서 초기 적용 시에는 무작위로 10문항을 추출하여 학습문제를 나타내며, 학습데이터가 누적될수록 오답율과 출제 빈도수 대비 적절한 문제를 추출하여 나타낼 수 있다.



<그림 10> Java학습문제 풀이

학습과정에 따른 문제풀이는 <그림 11>과 같으며, 이때의 학습 데이터베이스내의 추출된 문항번호는 <그림 12>에 나타내어 보았다.



<그림 11> 정답확인



<그림 12> 문제 데이터베이스의 문항번호

이러한 학습과정에 따라 <그림 13>에서 30회의 학습 시스템 적용 후에 대한 문제 데이터베이스의 내용으로서, 예제로 20번 항목까지 산출된 학습데이터의 내용을 나타내었다.



<그림 13> 30회 적용후의 예제 20문항에 대한 문제데이터베이스의 내용

즉, 오답율이 30%이하이며 출제율이 50%이상인 문항에 대해서만 미출제 처리가 되며 나머지는 모두 출제처리가 되어 문제 추출의 대상이 되어 그 중에서 무작위로 추출된다. 이는, <그림 14>와 같이 실제 출제되는 상황을 나타내었다.

문항번호	DB내의 출제문항번호
1	14
2	30
3	25
4	3
5	18
6	41
7	20
8	7
9	33
10	5

<그림 14> 30회 적용후의 출제문항

즉, <그림 13>에서 출제로 판정된 문항번호들이 <그림 14>에서 많이 출제됨을 알 수 있다.

5. 결 론

본 연구에서는 학습자의 학습효과를 높이고자 문항에 대한 오답율과 출제 빈도수를 고려하여 문제 데이터베이스에서 문제들을 추출하는 시스템을 구현하였다. 이에 따라, 학습자들이 학습성과에 따라 다양하게 변화되는 학습자료를 접할 수 있었다. 즉, 단순히 오답율만을 고려하지 않고 출제 빈도수까지 고려함으로써 자주 출제되면서 누구나 쉽게 풀 수 있는 문제들은 출제기준에서 제외시킬 수 있었다. 또한, 문제추출 알고리즘부분을 EJB 컴포넌트로 구현함으로써 중요 비즈니스 로직을 메인에서 분리시킴으로서 이를 별도로 처리, 구현, 관리할 수 있으며 유지보수 및 재사용이 용이하여 보다 효율적인 시스템 개발의 이루어질 수 있었다.

그러나, 본 연구는 출제유무에서 출제로 판명된 문항만을 우선적으로 추출함으로써 시스템의 사용이 증가될수록 출제보다 미출제 항목들이 많아지게 된다. 결과적으로 출제항목에서 추출하여야 할 문제들이 미출제 항목에서부터 다시 재적용 하여야 하는 불합리한 문제를 안고 있다. 따라서, 향후연구과제로는 위 문제를 해결하고 보다 많은 문항들을 적용하여 다양하고 효율적인

학습자료 제공이 이루어져야 할 것이다.

참 고 문 헌

- [1] 이충기, “웹 기반의 자바 프로그래밍 강의 교안 개발”, 공학교육연구 제5권 제1호, 한국공학교육기술학회, 2002.
- [2] 김민경외 4인, “아동의 문학 능력 신장을 위한 웹 기반 아동문학 교육활동 프로그램의 개발 및 효과검증”, 교육공학연구 제19권 제1호, 한국교육공학회, 2003.
- [3] 이철현, 구덕희, “인터넷 기반 문제해결 학습 모형 개발”, 한국정보교육학회 논문지 제6권 2호, 2002.
- [4] 이재무, “개인차를 고려한 웹 기반 코스웨어 개발”, 한국컴퓨터산업교육학회 논문지 제2권 제12호, 2001.
- [5] 김재생, “분산 컴퓨팅 환경에서의 웹 교육 컴포넌트 개발과정 모델링”, 한국정보교육학회 논문지 제6권 2호, 2002.
- [6] 최희석, 염근혁, “효과적인 소프트웨어 아키텍처 평가방법”, 정보과학회 논문지:소프트웨어 및 응용 제29권 제5호, 한국정보과학회, 2002.
- [7] 노규성, 최정웅, “효과적인 IT 교육용 e-Learning 컨텐츠 개발사례”, 정보처리학회 지 제9권 제5호, 한국정보처리학회, 2002.
- [8] 허원, “웹 기반의 자동 문제 출제 및 평가 시스템의 개발 및 활용”, 공학교육연구 제3권 제2호, 한국공학교육기술학회, 2000.
- [9] Alan W.Brown, “From Component Infrastructure To Component-Based Development”, Sterling Software, ICSE Workshop, 1998
- [10] 윤희진, 최병주, “컴포넌트 기반 개발 개념을 활용한 테스트 프로세스 tailoring”, 정보과학회 논문지:소프트웨어 및 응용 제27권 제12호, 한국정보과학회, 2000.
- [11] 김경아, 최은만, “웹기반 교육에서의 자동 문제출제 시스템”, 정보처리학회 논문지 A 제9-A권 제3호, 한국정보처리학회, 2002.



정 화 영

- 1991 목원대학교 수학교육과
(이학사)
- 1994 경희대학교 전자계산학과
(공학석사)
-
- 2000~2001 경희대학교 전자계산공학과
박사수료
- 1994~1998 아주시스템(주) 기술연구소
전임연구원
- 1999~1999 CNA Research(주) 전임연구원
- 2000~현재 예원대학교 정보경영학부
전자상거래학과 전임강사
2002. 12 (사)한국디지털컨텐츠학회
추계학술위원
2003. 5 (사)한국인터넷정보학회
춘계학술위원
- 2003~현재 (사)한국콘텐츠학회 학회지
편집위원
- 2003~현재 한국산업단지공단 전문 컨설턴트
관심분야: WBI, 웹 컴포넌트, CBD
- E-Mail: jhymichael@empal.com