

WBI상에서 XML요소를 이용한 교과 색인 시스템

채승천⁰, 배석찬
군산대학교

cc1489@hanmail.net, scbae@kunsan.ac.kr

Course Indexing system using XML Component in WBI

SeungChun Chae⁰, Seokchan Bae

Dept of Computer Education, Graduate School of Education KunSan University

요 약

오늘날 정보의 중요성의 증가와 신속한 정보검색의 필요성에 따라 정보관리의 현대화가 요구되고 있다. 이에 따라 색인생산의 전산화를 기하기 위한 시스템을 활용하지 않은 전자문서의 표준화로 기술된 XML의 요소를 정보검색을 할 수 있도록 설계를 통해 일반 웹 문서가 전자문서로 서로 교환될 수 있다. 그리고 색인기능을 지원하여 정보를 검색하고 기존의 텍스트와 여러 다양한 형태의 정보를 구조적으로 표현할 수 있어 웹을 통한 학습자에게 웹 문서를 쉽게 습득할 수 있다. 본 논문에서는 구조화된 문서의 표준인 XML의 요소를 이용한 교과색인 시스템을 웹에서 이루어지는 형태의 학습자료로 보다 쉽게 만들고 문서의 교환과 내용에 대한 색인이 원활히 이루어져 교사와 학생간의 상호작용하는 교과색인 시스템을 설계 및 구현하였다.

1. 서 론

최근 인터넷의 발전과 더불어 인터넷을 사용하는 사람들이 많이 늘어났다. 또한 웹 활용 교육에 대한 관심이 증가해가면서 원격강의 또는 실내에서 컴퓨터를 통한 교육을 실시하기 위해 많은 노력을 하고 있다. 특히 웹을 통해서 학습을 하기 위해서는 사용자들은 웹 문서의 내용을 파악하거나 문서 내에서 필요한 부분을 찾아가야만 한다. 인터넷을 이용하여 교육할 경우 시간과 공간이라는 제한을 벗어나 여러 가지로 표현된 형태의 정보를 활용하면 학습자에게 보다 넓은 지식과 기능을 발휘할 수 있도록 한다. 또한 인터넷을 이용하면 학습자들은 이미 지나 그림을 통해서 이루어지기 때문에 현재 종이로 만들어진 교과보다는 흥미와 동기유발에 큰 도움이 될 것이다. 기존의 문서가 단순 텍스트로 생산되고 문서 생산에서 정보를 추출한다는 것은 몇 가지 제한이 뒤따르게 된다. 첫째, 문서의 양적인 면에서 볼 때, 문서가 많아지면 문서를 검색하는 데 걸리는 시간은 문서의 양에 비례하게 된다. 둘째, 문서의 질적인 면에서 볼 때, 한 문서 내에서도 사용자가 필요한 정보가 들어있는지 없는지는 일단 검색을 해보아야 한다. 그러므로 웹 문서가 만들어지면 문서는 검색하는 색인기능이 필요하므로 색인기능은 확대되어야 한다.[1][9]

본 논문에서는 색인에 필요한 구조적 문서에 필요한 요소를 제공하기 위한 XML의 요소를 적용하여 2장에서는 구조적 문서에 대한 관련연구 고찰로 웹 상에서 실행

되는 교과의 기능을 제시하였다.

3장에서는 본 논문에서 제안한 방법을 설명하였다. 현재 웹 문서의 HTML문서와 XML언어로 구성된 구조적 요소를 사용, 웹 상에서 교과 색인을 구현한다. 4장은 본 논문의 결론과 향후 연구방향을 기술하였다.

2. 기존 연구

기존의 데이터베이스 시스템이나 현재까지의 관습적인 정보검색 시스템들에서는 문서를 연속된 단어들의 조합으로 인식하여 검색 시에 문서를 검색간 위로 간주하거나 전문검색하고 또한 일반 문서는 단순 텍스트 외에 문서의 장, 절, 제목, 참고문헌 등 다양한 논리적 구조정보를 가지고 있다. 이러한 구조정보는 사용자가 전체문서의 내용을 파악하거나 문서 내에서 필요한 부분을 찾아가는데 매우 중요한 역할을 한다[8]. 이러한 구조적 문서를 표현하기에 알맞은 SGML 이 거의 사용하지 않은 불필요한 기능들을 삭제한 XML의 요소를 사용하여 문서 내에 문서가 가지는 구조정보를 명시적으로 가지는 문서를 구조적 문서(structured document)라고 한다. SGML(Standard Generalized Markup Language)이나 XML(eXtensible Markup Language)은 문서가 가지는 구조를 마크업으로 구분된 요소를 사용하여 명시적으로 기술하는 언어이다. 또한 웹 상에서 구조화된 문서를 전송 가능하도록 설계한 마크업 언어이다. 현재 웹의 확산으로 인터넷의 사용이 증가하면서, 웹에 XML을 이용하

여 구조화된 문서를 작성할 수 있는 다양한 XML을 지원 하는 소프트웨어들이 개발되고 있다. 앞으로 XML문서가 웹에서 확산될 전망으로 XML로 작성된 문서에 내장된 구조에 대한 정보를 활용한 검색 시스템 개발이 필요하다. 이러한 장점을 살려 교사-학생간의 상호 작용하는 매체중의 하나인 교과서 색인 시스템에 적용한다.

HTML은 기본적으로 VIEW에 입각해 있기 때문에 효율적이고 통일적인 전자문서 또는 데이터관리에는 역부족임에 틀림없다.[2] 그러므로 본 논문에서는 SGML과 HTML, XML의 비교로 구조화문서를 웹에서 표현할 WBI기반과 현재 웹 강의 분석을 비교하였다.[5]

2.1 구조화 문서

구조화 문서는 문서 내에 내용과 함께 문서의 구조를 포함하고 있는 문서임 구조는 다양한 논리적인 정보와 관계들을 나타낸다. 1986년 SGML이 전자문서의 전송, 저장 및 자동처리에 대하여 ISO(International Standard organization) 표준으로 채택되었으며 [ISO8879] 최근의 문서들이 이와 같은 구조화된 형식으로 만들어지는 추세이다[7]. 최근에는 구조화된 전자문서가 급속히 증가함에 따라 SGML과 XML같이 구조화된 문서를 체계적으로 공유, 저장, 검색해야 할 필요성과 이에 대한 연구의 중요성이 증대되고 있다[7][9]

2.2 XML이 HTML이나 SGML보다 앞선 기능[4]

- ① Tag의 종류와 표현력을 향상시키기 위해 HTML과 달리 대소문자가 구별된다.
- ② SGML과 달리 DTD 없이도 문서 제작과 배포가 가능하다.
- ③ 모든 Tag는 시작과 종료의 쌍으로 이루어진다.
- ④ 수식 표현이 가능하고, 양방향링크, 다 방향 링크, Spanning 링크 등 보다 발전된 문서간의 링크를 제공한다.
- ⑤ 링크자체에 대한 정보를 검색이나 취득의 대상으로 삼을 수 있다. 즉 의미를 지닌 링크가 가능하며 이런 링크의 의미를 대상으로 한 검색이 가능하다.

국내에서 구축되는 WBI에 의한 원격 수업의 형태는 <그림 1>과 같다



<그림 1> 원격 수업의 절차

즉, 교사와의 상호관계가 아닌 고정적인 교과 내용과 한정된 범위에서의 수업형태로서 학생들은 이미 만들어진 동영상강의와 요점정리 및 평가를 통해 교과 내용을 습득하고 있다. 내용면에서도 많은 습득보다는 교육자의 모습을 통한 강의를 주로 하고 학생들의 지나간 수업에 대한 복습은 잘 이루어지지 않은 실정이다. 이러한 경우

색인할 수 있는 환경이 구축된다면 학습과 복습에 대한 효율성이 증대될 것이다.

3. WBI에 기반을 둔 XML요소를 이용한 교과색인 시스템 설계 및 구현

교과서의 내용이 입력된 XML파일을 로그 하기 위해 <그림 2>처럼 CreateObject를 사용하여 변수를 선언하여 DOM 오브젝트에 접근한다.[3][4]

```
set xmlDoc=createObject(Microsoft.XMLDOM")
xmlDoc.async=false
xmlDoc.load("book.xml")
```

<그림 2> 로드할 문서 액세스

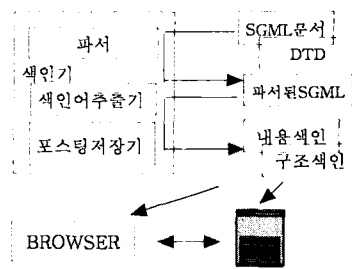
<그림 3>처럼 외부 텍스트를 위해 또는 링크를 위한 텍스트노드를 지정하여 텍스트를 꺼낸다.

```
<xsl:template match="text()">
<xslvalue-of />
</xsl:template>
```

<그림 3> 링크 요소

3.1 SGML 방법 분석

SGML로 작성된 문서는 HTML의 효시인 동시에 하이퍼미디어 텍스트의 DTD를 적용한 웹언어이다. 하지만 웹 개발을 위한 목적으로 만들어진 언어가 아니다. 더욱이 ISO8879에 정의된 소프트웨어산업 전반 표준으로써 문서의 구조와 내용을 기술한 것이다. 즉 구조체언어의 일종이다. 그러므로 다양한 장점과 다양한 기능을 가지고 있음에도 불구하고 사용이 어려워져 <그림 4>처럼 SGML의 문서는 DTD를 동반해 내용색인과 구조색인이 동시에 이루어져야만 한다[7].



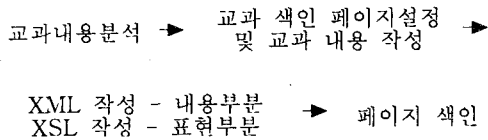
<그림 4> SGML 문서의 브라우징 형태

결과적으로 SGML원문서와 해당 DTD를 파서를 통해 분석한다. 문서의 전처리가 필요한 이유는 SGML문서의 DTD와 SGML문서의 일치여부를 확인해야 하기 때문이다. 문헌의 논리적 구조를 기술하기 위한 메타언어인 SGML로 만들어진 문헌은 크게 SGML선언부(SGML Declaration) 문헌 유형 정의부(Document Type Definition: DTD)그리고 문헌부 (Document Instance: DI)의 세 부분으로 나눌 수 있다.[9] 이 가운데 DTD는

SGML의 가장 핵심적인 부분으로서 문헌에 나타나는 논리적인 구조를 설명하는 문법으로 구성된다. 특히 문헌의 구조적 특성의 표현, 특히 문헌의 구조적 특성의 표현 즉, 계층적 문헌 구조를 파악해보는데 적절하다. 그러므로 다른 유형에 대한 문헌의 분석도 DTD에 달려 있다고 말할 수 있다[6]. 본 논문에서는 SGML의 한 부분인 문헌유형 정의부의 구조에 대한 규정이 제한되지 않은 XML 문서중 정형식(Well Formed)문서의 자율성에 맞고 복잡하지 않는 XML로 작성된 문서의 요소와 속성을 이용한 색인 기법을 학습할 수 있는 전자교과서 색인시스템을 연구하였다. 또한 XML문서는 내용과 표현이 분리되어 내용을 데이터베이스화하고 표현은 관리자나 교육자가 스스로 변경하여 교과서의 흥미와 색인의 형태를 빠르게 바꿀 수 있다.

3.2 시스템 구성도

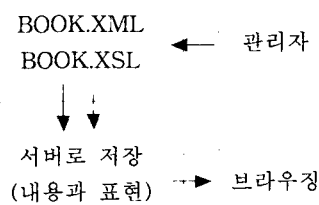
본 논문은 전자교과서의 전체 시스템에 활용할 수 있는 색인시스템이므로 색인과정만을 구성한다. <그림 5>에서와 같이 전체 시스템 구성도를 기술하였다.



<그림 5> 시스템 구성도

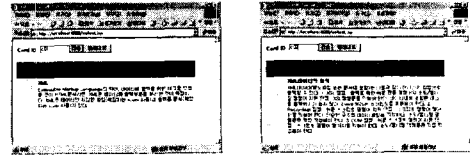
3.3 학습 흐름도 및 색인 모듈

본 논문은 서버와의 관계부분은 제시하지 않았으므로 학습의 색인부분을 화면에서 다룬다. <그림 6>은 <그림 5>의 색인시의 모듈을 나타내었다. 교사(관리자)는 교과 색인을 더욱 시각화하기 위해 표현부분을 학습 전에 준비하여 학습도중이나 학습 후에 색인시스템을 이용하여 학습이 원활히 이루어질 수 있도록 한다. 학생(학습자)는 색인아이콘을 눌러 입력 폼에 찾고자 하는 색인 페이지를 입력 해당 페이지를 찾아갈 수 있다. XML 교과서로 펼쳐지면 교과서의 목차가 생겨 어떤 페이지로 가서 텍스트를 쉽게 찾아 갈 수 있다.



<그림 6> 색인 시 모듈

따라서 전자교과서의 중요기능이 되어야 할 색인기능으로서 본 논문에서는 XML문서의 기본 형태를 가지고 XML의 요소를 인덱스 코드를 이용하여 색인시스템을 구현하였다.



<그림 7> 페이지 색인 환경

4. 결론 및 향후연구방향

본 논문은 교육에 필요한 교과를 웹 상에서 문서교환과 작성이 쉬운 XML요소를 이용하였다. 이처럼 여러 사용자들의 연결로 시스템의 부하를 극복하고 XSL을 이용, 전달되는 정보를 텍스트를 구조화하여 빠른 색인과 표현의 여러 가지 방법으로 시각화할 수 있다.

아울러 흑판과 분필위주의 강의식 수업에서 다양한 교육방법을 적용하는 열린 교육으로 나아가야 한다. 그러기 위해서는 멀티미디어를 활용하기 위하여 보다 빠른 학습을 위한 색인 정보와 그에 따른 방법으로 수업 설계의 편의(정보전달과 판서의 구조화)를 제공하는 데는 멀티미디어 자료를 활용하는 것이 필요하다고 판단된다.

앞으로 웹을 통한 교육상뿐만 아니라 교육정보화에 필요한 색인을 좀더 여러 분야에 사용되기 위해서는 XML과 같은 구조적 문서에 알맞고 보다 쉬운 체제와 정적이 아닌 그림과 동영상을 추가한 동적인 웹 효과를 표현할 수 있는 기법이 연구되어야 할 것이다.

참고문헌

- [1] 김정은, 자동색인의 색인어 선정기법에 나타난 색인의 의미 전남대학교 석사학위 논문 1995. 8.
- [2] 박정훈, XML을 이용한 문제출제 시스템 구현 금오공과대학 석사학위 2000. 5
- [3] 원두영, 이차미 표준 XML 데이터전달과 저장을 위한 차세대 인터넷 언어 1998
- [4] 한국 컴퓨터매거진, 세계에서 가장 쉬운 XML, 2001
- [5] 최정숙, WBI에 기반한 초등학교 컴퓨터 교육용 전자교과서 설계 및 구현 1998. 12
- [6] 이재진, 객체지향데이터베이스를 이용한 구조화된 문헌의 저장 및 관리에 관한 연구 연세대학교 석사학위논문
- [7] 장현철, 관계데이터베이스를 이용한 구조화 문서의 색인 및 검색 충남대학교 석사학위 논문 1999. 2.
- [8] 김영자, 구조적 문서의 효율적 검색을 위한 색인 색인 구조 경상대학교 석사학위 논문 2000. 8.
- [9] 유석중, SGML한글 문서의 논리적 구조에 근거한 자동색인기법에 관한 연구 연세대학교 석사학위논문 2000.