

## 멀티미디어를 활용한 가상대학 XML 문서에서 Paser의 역할

공범용\* 황병곤\*\* 조세홍\*\*\*

\* 영남이공대학 \* \* 대구대학교 \* \* \* 한성대학교

### The Role of Paser for XML Documents using Multimedia in the Cyber University

Beom-Yong Kong\*, Byung-Kon Hwang\*\* Sae-Hong Cho\*\*\*

\* Yeungnam College of Science & Technology  
Email: bykong3@ync.ac.kr

\*\* School of Computer Information Engineering, Daegu University  
Email: bkhwang@taegu.ac.kr

\*\*\* School of Computer Engineering, Hansung University  
Email: chosh@hansung.ac.kr

#### 요 약

최근 인터넷기술이 발전하면서 교육환경에서 멀티미디어 기술을 활용한 패러다임 전환이 이루어지고 있고, 이 전환에 가상대학에서 사용되어지고 있는 문서의 역할의 중요성은 크게 주목을 받고 있다. 문서의 표준화를 위하여 구조적인 문서 처리가 가능한 XML이 대두될 것이다. XML에서 논리적 구조를 표현하는 방법은 사용자 정의인 문서타입(DTD)과 XML Schema의 두 가지 방식이 있으며, 향후 멀티미디어를 활용한 XML 문서에서 Paser의 역할이 중요하게 될 것으로 예측되어진다. 이에 따라 본 논문에서는 멀티미디어를 활용한 XML 문서에 필요한 자료들을 분석하고 가상대학에서 발생하는 문서행정업무를 효율적으로 지원하는 방안을 제안하는 XML 문서에서 Paser의 역할을 제시하고자 한다.

#### 1. 서 론

인터넷을 이용한 원격교육이 몇 년 사이에 급 성장되고 있다. 본 논문에서는 가상대학의 학사 운영 및 관리와 관련하여 다양한 이슈들 중 멀티미디어를 활용한 가상대학 XML 문서관리시스템에서 Parser의 역할을 제시하고자 한다. 가상대학 뿐만아니라 Web

상에서 멀티미디어 문서가 전달될 때, 문서의 종류와 업무량이 증가할 것임으로 사용자가 요구하는 XML 문서를 보다 효과적으로 이용할 수 있는 기술이 또한 요구된다. 이러한 요구에 부응하고자 등장한 것이 바로 멀티미디어를 활용한 XML(eXtended Markup Language)로서, 1996년 표준제안 이후 현재까지 XML 문서의 기능이 확장되

고 있는 상황이다[1].

본 연구는 인터넷 중심의 정보교환 사회에서 멀티미디어를 활용한 XML문서에서 행정업무의 효율성을 제고하기 위한 문서처리 수반을 공유하기 위한 인터넷 XML 문서에서 Paser의 역할에 대한 방안을 제안 하고자 하였다.

## 2. 관련기술 및 응용

### 2.1 XML 응용기술

XML(eXtensible Markup Language)은 기존의 HTML이 갖는 장점인 웹 상에서 정보제공과 SGML이 갖는 장점을 이용하여 다양한 구조의 문서제공에 활용할 수 있도록 보완한 차세대 웹 언어의 표준이다. 따라서 고정된 형식의 언어가 아닌 확장 가능한 언어로서 사용자의 필요에 따라 정의하여 사용할 수 있으며 문서구조는 다음과 같이 물리적 구조와 논리적 구조로 구성되어 있다[2].

첫째 XML 문서의 요소

- 엘리먼트(Element) : 논리적 구조 표현의 기본단위
- 엔티티(Entity):XML문서에서 Paser의 처리를 요하는 문서의 최소 단위다.
- 속성(Attribute) : 엘리먼트나 엔티티의 추가정보를 기술하기 위한 단위다.

둘째 XML 문서의 구성

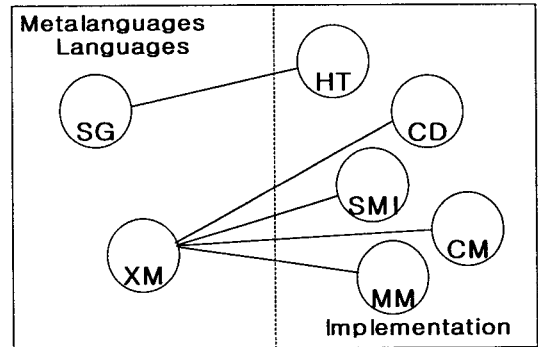
- XML 문서형 정의(DTD: Document Type Definition)
- XML 문서의 논리적 구조를 정의
- XML 문서실례(XML Document Instance).

셋째 문서형 정의에 따라 작성된 XML 문서이다. 그러나 기본적인 구조는 HTML을 포함하여 다른 대부분의 SGML 어플리케이션과 유사하다. 그러나 XML 문서는 문서

타입 선언(document type declaration)없이 저자 자신이 설계한 마크업 안에 문서를 집어넣어 간단하게 작성될 수도 있다.

### 2.2 XML 활용기술

디스플레이를 위해 데이터를 포맷하는 방법을 기술하는 HTML과 달리 XML은 데이터 구조의 의미(semantics)를 정의하고, 다른 마크업 언어를 만들기 위한 메타 언어로 사용된다.



<그림 1> 메타데이터로서의 XML

XML은 마크업 언어의 원조적인 SGML의 기술을 많이 채용하고 있다[3]. 각 SGML 파일들은 적용 가능한 태그와 태그의 사용규칙을 포함하는 DTD를 포함하며 문서의 내용을 확인하기 위해 파서들을 사용한다. XML은 SGML이 가지는 기능의 많은 부분을 제공하면서도 관련 소프트웨어 개발이나 웹 상의 구현이 용이하다. XML은 SGML과 같이 태그집합들로 구성되는데 태그들은 그 자체로는 메타데이터이며 문서구조, 의미(semantics), 스타일 등에 관한 정보를 제공한다[4]. 태그는 어떠한 이름도 가능한데, 이것이 XML에 확장성을 부여한다. DTD는 특정문서나 애플리케이션에서 태그를 정의하고 해석하는 규칙을 제공하는 것으로서XML에서는 선택적이다[5]. DTD는 XML 문서속에

있을 수도 있고 외부로부터 링크 될 수도 있다. DTD를 포함하지 않은 문서는 well-formedness[6]가 확인되어야 하고, DTD를 포함하는 XML문서는 DTD규칙을 잘 따르는지가 확인되어야 한다.

### 3. XML문서에서 Paser 및 그 역할

#### 3.1 XML Paser

병용처리모듈 XML Paser는 XML 파일이 이미 존재하는 경우에는 파일을 입력받아 병용처리모듈을 통해 문서를 구분하고 파서를 통해 DTD 정보와 실제 문서부 정보를 구분하여 엘리먼트, 어트리뷰트, 엔티티 등의 문서논리 구조 정보인 DTD 정보를 DTD 제너레이터로 보내고, 실제 문서부 정보를 Document Instance Editor로 보낸다.

실제 문서부 편집기에서는 DTD에서 정의해 놓은 문서구조에 맞추어 XML문서를 작성할 수 있다. DTD 생성기를 통해 DTD파일을 편집할 수 있으며 편집된 DTD 정보에 따라 실제 문서부 편집기에서는 XML 문서를 편집할 수 있다. 작성된 문서를 프리젠테이션하기 위해서는 XSL 생성기에서 작성된 XSL파일정보를 DTD와 결합된 실제 문서부 정보가 입력받아 포맷을 통해 프리젠테이션된다.

#### 3.2 Paser의 역할

HTML 문서들을 잘 정리되게 즉, well-formed 하게 만든다. XML에서 DTD는 선택사항이다. 그러나 HTML 파일들은 HTML DTD의 XML 버전이 아직 없기 때문에 현재 어디에서도 DTD가 없을 수 밖에 없다. 대부분의 HTML DTD들은 마침 태그

의 생략을 허용하지만 XML의 경우는 그러하지 않기 때문이다. 이미 많은 HTML 저작 도구들은 아주 충분하지는 않지만 거의 well-formed한 XML을 만들어낸다.

○ 모든 XML 문서들은 well-formed하여야 하지만(아래를 참조하라), DTD는 조건부이다. HTML 파일들은 XML의 DTD가 없는 형식으로 변환될 수 있다. 그러나 현재 SGML의 HTML DTD들의 XML 버전은 있을 수 없다.

○ 만약 여러 개의 HTML Document Type Definitions (DTDs) 중 하나를 따르는 HTML 파일을 만들었고 그것들이 유효하다면(validate OK), 그것들은 다음과 같이 XML로 변환될 수 있다.

즉 DOCTYPE 선언(declaration)과 내부의 부속항목들(internal subset)을 확인한다. 이는 기본적으로 <!DOCTYPE HTML...>의 격쇠 안에 들어 있는 모든 것들이 해당된다 --을<?XMLversion="1.0"standalone="yes"?>과 같은 XML 선언으로 바꾼다.

○ 모든 독립요소(EMPTY elements)들의 오른쪽 격쇠를 '/'>'한다.--예를 들면 <IMG SRC="mypic.gif" alt="Picture"/>로 바꾼다. 독립요소들로는 다음과 같은 것들이 있다. 문서의 헤더(header)에서 사용되는

<ISINDEX>, <BASE>, <META>, <LINK>, <NEXTID>, <RANGE>와 문서의 바디(body)에서 사용되는 <IMG>, <BR>, <HR>, <FRAME>, <WBR>, <BASEFONT>, <SPACER>, <AUDIOSCOPE>, <AREA>, <PARAM>, <KEYGEN>, <COL>, <LIMITTEXT>, <SPOT>, <TAB>, <OVER>, <RIGHT>, <LEFT>, <CHOOSE>, <ATOP>, <OF>가 있다.

즉 마침 태그가 필요한 요소들(non-empty elements)이 정확히 사용되었는지 확인한다. 예를 들어 <P>는 항상 </P>로 닫혀져야 한다. 이 작업은 SP의 일 부분인 sgmlnorm과 같은 표준화된 프로그램이나 Emacs/psgml과 같은 에디터의 sgml-normalize 기능에 의해 자동적으로 처리 될 수 있다.

○ 마크업에 사용되는 문자인 <과 &를 &lt;와 &amp;로 바꾼다.

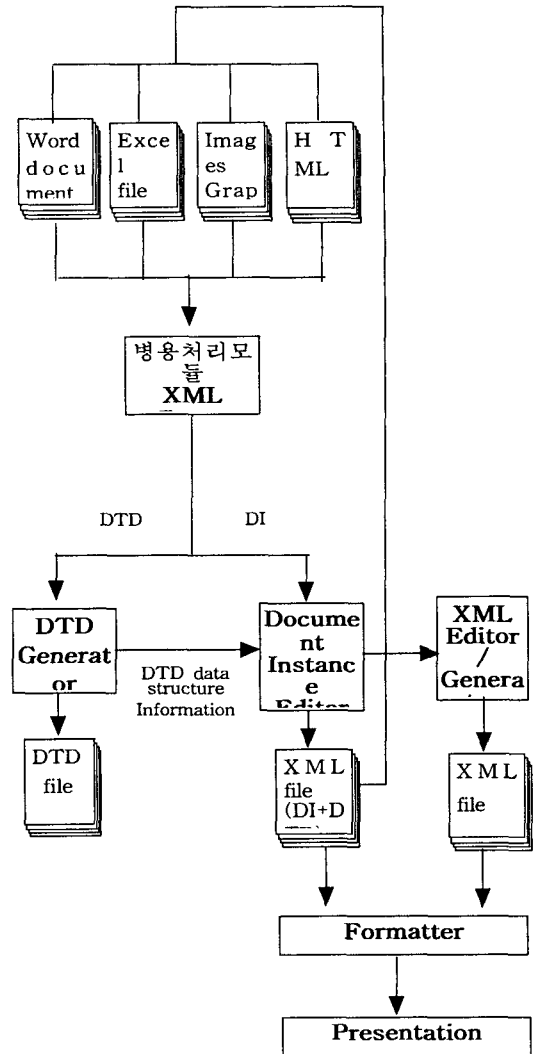
○ 모든 속성값(attribute values)이 따옴표 안에 있는지 확인한다.

○ 요소의 이름들이 나올 때 시작태그와 마침태그에서 대소문자 구별이 정확히 일치하는지, 그리고 전체문서에서 일관적으로 사용되는지 확인한다.

○ 모든 속성의 이름들이 전체문서를 통해 일관적으로 사용되는지 확인한다.

HTML 브라우저들은 XML 스타일의 독립 요소(EMPTY elements)들의 종료처리(trailing slash)를 이해하지 못하기 때문에 위와 같이 변환된 것은 반대로는 호환되지 않는다. 하나의 대안은 <IMG>를 <IMG></IMG>로 바꾸는 것과 같이 모든 독립요소에 가짜 마침 태그를 붙이는 것이다.

만약 많은 양의 유효한(valid) HTML 파일들을 가지고 있다면, 이러한 작업을 SGML 변환 시스템--Omnimark, Balise, SGMLC, 또는 SGML Perl라이브러리 사용하는 시스템 등이 있다--에서 스크립트를 이용하여 처리할 수도 있으며, 하고자 하는 작업을 잘 알고 있다면 에디터의 매크로 기능을 이용해서 <그림 2>와 같이 처리할 수 있다.



<그림 2 > Paser의 역할

#### 4. 결론

인터넷 기술이 발달되면서 원격교육에 대한 관심도 증폭되고 있다. 본 논문에서는 아직까지 크게 활성화되지 못하고 있는 가상대학 원격 교육활성화를 위한 XML문서가 가상대학의 문서 시스템의 표준이 될 것으로 예측하고, 이 문서들의 효율적인 관리를 위하여 XML Paser의 역할을 제안하였다.

XML형태의 문서를 효율적으로 저장, 관

리, 검색할 수 있도록 하는 Mechanism에 대한 고찰은 가상대학의 성패를 가름할 수 있는 중요한 요소인만큼 앞으로 꾸준한 연구가 지속되어야 할 것이다.

### 참고문헌

- [1] Buford, J. & Koegel, F., *Multimedia Systems*, Addison Wesley Publishing Co., 1994.
- [2] 정희경."XML가이드" 그린 1998.
- [3] 최 영, 『미디어시대의 네트워크 커뮤니티 케이션』, 박영출판사, 1998, pp.12~13.
- [4] Baldwin, T.F., McVoy, D.S., and Steinfield, C., *Convergence: Integrating Media, Information & Communication* (California; Sage Publications, Inc., 1996), pp. 3 - 4.
- [5] 최지태.전영훈.황대훈."DTD생성기능을 지원하는 XML문서 병용 편집기의 설계 및 구현 ".한국멀티미디어추계논문집. 1999.
- [6] 김동신.이용규."XML 문서접근 권한관리 ".한국멀티미디어 추계논문집.1999.