

XML 을 이용한 논문 참고문헌 자동 삽입 처리기의 설계 및 구현

김병규⁰, 강무영*, 박재원*, 강지훈**
한국과학기술정보연구원, 충남대학교 컴퓨터학과**
{yourovin,kmy,ilonetos}@kisti.re.kr, jhkang@cs.cnu.ac.kr

Design and Implementation of Auto Insert Processor For Reference of Paper using XML

Byung-Kyu Kim⁰ Mu-Yeong Kang* Jae-Won Park* Ji-Hoon Kang**
Korea Institute of Science and Technology Information*
Dept. of Computer Science, Chungnam National University**

요 약

논문을 작성하기 위해서는 참고문헌을 달아주어야 하는 쉽지 않은 과정이 있다. 일반 사용자가 쉽게 작성하고 적은 노력으로 참고문헌을 삽입시켜줄 수 있는 처리기가 필요하다. 기존의 LATEX 시스템은 고정된 환경과 많은 제약사항으로 사용하기에 불편이 많았다. 본 논문에서는 XML의 구조적인 특성을 이용하여 참고문헌을 자동으로 삽입하는 새로운 어플리케이션을 제안한다. 이를 위해 논문과 참고문헌 사전에 대한 문서타입을 정의하였으며 참고문헌사전의 구조 정보를 검색, 삽입, 수정, 삭제 할 수 있는 사전 관리기를 설계하였다. 이를 바탕으로 참고문헌을 자동 삽입 처리된 논문 끝 XML문서는 논문 구조에 맞게 정의된 XSL을 통해 HTML로 변환되며 자유로운 웹 게시가 가능해진다.

1. 서론

XML(eXtensible Markup Language)은 1998년 W3C(World Wide Web Consortium)에 의해서 표준으로 채택된 전자문서의 기술을 위한 메타 언어(Meta Language)이다. XML은 SGML(Standard Generalized Markup Language)의 복잡한 면을 제거하고 구조, 검증, 확장의 특징을 계승하면서 현재 웹 문서 기술 언어로서 널리 사용되고 있는 HTML의 단점을 보완하여 차세대 웹 문서의 표준으로 자리잡고 있다[1,2].

XML 문서는 문서의 구성 요소인 엘리먼트(element)간의 논리적인 계층구조를 가지고 있는 DTD(Document Type Definition)와 DTD에 선언된 엘리먼트에 따라 기술된 XML 문서내용(Document Instance)을 가진다. XML 파서(parser)는 DTD를 이용하여 XML문서의 유효성 검사를 하며 이것은 XML 문서 안에 기술된 구조화 된 내용이 DTD에 선언된 엘리먼트 타입의 논리적 계층구조와 같은 지를 검사하는 것이다. 그리고 유효한 XML 문서내용은 XML 응용 프로그램에 제공되어 사용되어진다. 최근 XML 문서를 다양한 분야에 적용하기 위해서 XML 응용 프로그램에 대한 많은 연구가 되고 있다.

본 논문에서는 이러한 XML의 특징을 이용하여 논문 작성 시에 LaTeX System에서 지원하는 참고문헌 자동 삽입 기능과 참고문헌 사전관리를 효율적으로 할 수 있는 처리 시스템을 설계하고 구현하였다.

논문의 구성은 먼저 3절에서 xml을 이용한 참고문헌 자동 삽입처리의 기본 개념을, 4절에서는 자동 삽입 처리 시스템의 설계 및 구현을 설명하고 5절에서 관련연구 내용을 살펴며, 마지막으로 5절에서 결론을 맺는다.

2. XML을 이용한 참고문헌 자동 삽입 처리

참고문헌 자동 삽입 처리란 논문을 작성할 때 자신이 인용한 참고문헌을 찾아 인용한 부분에 침식을 넣고 본문 뒤에 해당되는 참고문헌을 붙여 넣는 일을 자동으로 처리한다는 의미이다. 논문의 작성은 많은 수정의 반복이 필요한 까다로운 작업인데 논문 내용이 변경 될 때마다 참고문헌 삽입 처리의 과정을 매번 다시 해주어야 하는 작업은 작성자에게 많은 인내심을 요구 하며 불필요한 시간의 낭비를 초래한다. 그리고 사람의 수동적인 작업은 자칫 오류를 발생할 가능성이 크며 가장 중요한 문제점은 새로운 논문을 작성할 때마다 이 같은 작업을 반복해야 한다는 것이다.

이 때문에 개정이 가능한 주제별 참고문헌 사전이 필요하다. 사전이 갖는 각각의 참고문헌들이 고유한 키를 하나씩 가지고 있다면 논문 작성자가 참고문헌 삽입을 원하는 위치에 이 키만 적절히 사용해서 불편한 작업을 거치지 않고 원하는 참고문헌을 식별하고 삽입 해줄 수 있을 것이다.

본 논문에서는 위와 같은 “참고문헌 자동 삽입 처리”를 위해

XML을 이용한다. XML에서는 HTML문서에서 화면에 표현되는 형태만 지정하여 태그의 내용에 대한 아무런 정보도 제공하지 못하는 한계를 극복하기 위하여 문서 작성자가 자유롭게 태그를 생성할 수 있도록 확장되었다.

이렇게 구조화된 XML문서는 DOM 파서를 이용하여 태그의 종류별로 검색, 삽입, 삭제, 수정이 가능해 질 수 있다[3].

이러한 XML의 특징은 참고문헌을 자동으로 삽입시켜 줄 수 있는 배경이 된다.

3. 참고문헌 자동 삽입 처리 시스템 설계 및 구현

3.1 시스템의 구성

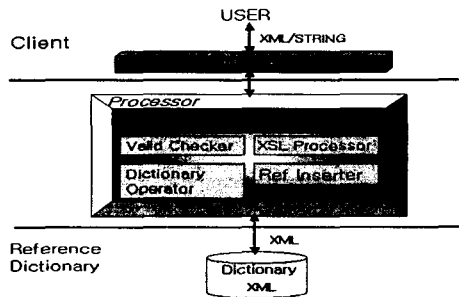
참고문헌 자동 삽입 처리 시스템의 구성은 다음 [그림]과 같다. 시스템은 4가지의 모듈로 이루어지며 각각의 기능을 살펴보면 다음과 같다.

모듈1): 논문 문서를 읽어 들인 후 유효성을 검사할 수 있는 기능을 담당한다.

모듈2): 논문에서 참고한 부분에 표식(알파벳 or 숫자)을 달아 주며 논문 문서의 끝에 해당 참고문헌을 삽입시켜준다.

모듈3): 완성된 논문.xml 문서와 XSL을 적용하여 볼 수 있도록 하는 기능을 수행한다[4].

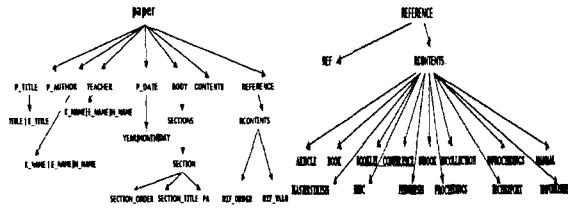
모듈4): 참고문헌을 종합적으로 관리 해 주는 역할을 한다.



(그림 1) 시스템의 구성

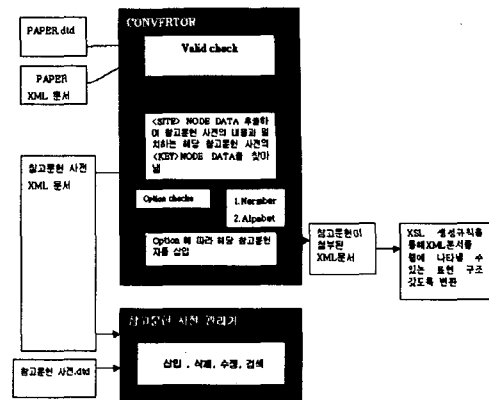
3.2 동작 원리 및 구현

논문과 참고문헌에 대한 문서정의는 그림 2와 같다.



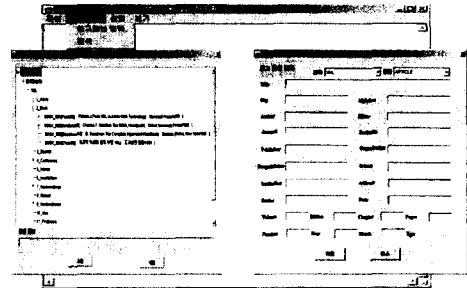
(그림 2) 논문과 참고문헌 사전의 DTD 구조도

위 문서구조를 바탕으로 한 논문에 대한 참고문헌 자동 삽입 처리기에 대한 개괄적인 동작 흐름을 살펴보면 그림 3과 같다.



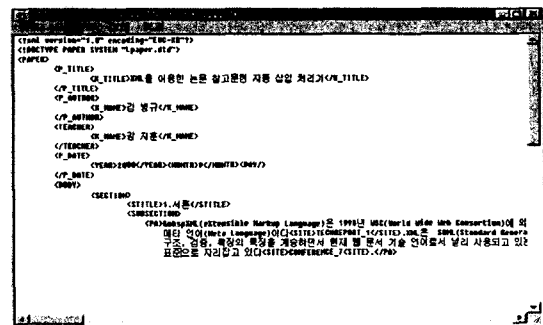
(그림 3) 동작 흐름을 나타냄

그림 4는 참고문헌 사전 관리기의 인터페이스이다.

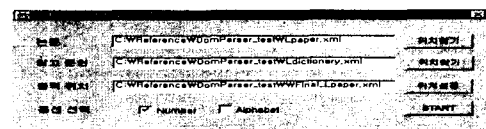


(그림 4)참고문헌사전 관리를 위한 화면

그림 5와 같이 참고한 부분이 마크되어 있는 논문에 대한 참고문헌의 자동삽입을 위하여 그림 6의 사용자의 옵션 선택(알파벳 or 숫자)을 사용한다.

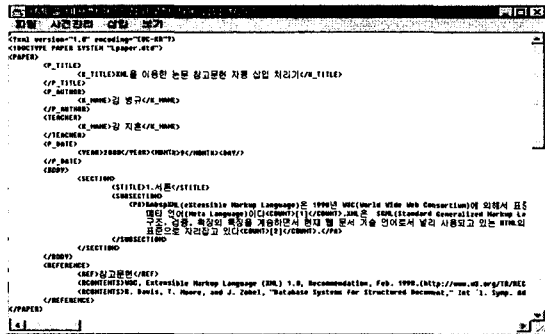


(그림 5) XML로 작성된 논문

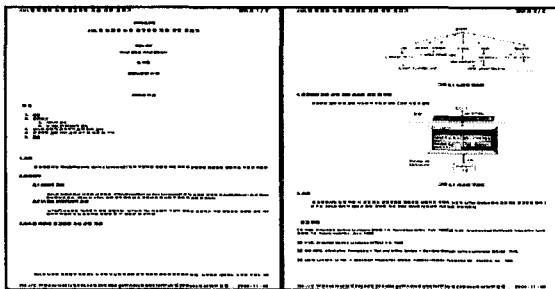


(그림 6)자동 삽입 처리 옵션

논문의 엘리먼트<SITE>의 키(TECHREPORT_1,CONFERENCE_7)를 엘리먼트<COUNT>로 바꾸어주며 COUNT의 값은 SITE의 개수 만큼 증가되는 수가 된다. 그리고 SITE의 키를 참고문헌 사전과 비교하여 일치하는 해당 참고문헌을 논문에 삽입시킨다.(옵션이 ALPHABET으로 선택된다면 COUNT의 값을 알파벳 생성규칙에 따라 만들어진 표식으로 한다.) 그 결과는 그림 7과 같으며 마지막으로 논문 구조에 맞게 정의된 XSL로 변환된 결과는 그림 8과 같다.



(그림 7) 참고문헌이 삽입된 논문



(그림 8) XSL로 처리된 논문

4. 관련연구

2.1 LaTeX System의 정의와 특징

LaTeX[1]시스템은 Tex의 특정 버전이다[5]. LaTeX가 Tex 시스템의 수학적 문법을 표현하기 위한 방안으로 개발된 것에 사용자 구분과 포매팅 명령어를 추가한 것으로 문서의 작성에 보다 효율적인 편집을 할 수 있도록 한 것이다.

이러한 LaTeX의 특징을 살펴보면 다음과 같다.

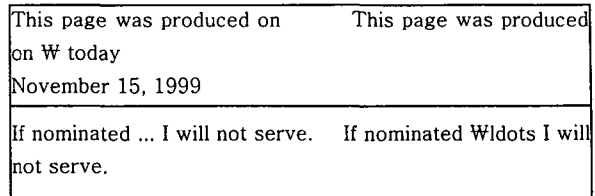
첫째로 문자의 치환이다. 문자의 치환은 특정 문자 또는 문자열을 한번에 다른 문자 또는 문자열로 변경해 주는 것을 의미하며, 심지어 .(마침표)다음에 스페이스로 치환도 가능하다.

둘째로 문자 생성 명령어로 문자열을 변수와 같이 정의해 놓고 사용하는 것이다. 이러한 예는 예 1과 같다.

셋째로 문서 형태를 지정하는 것이다. 이러한 문자의 형태 지정은 `\begin{document}` 형식을 취한다. LaTeX문서의 예는

예 2와 같으며 이렇게 지정된 문서는 LaTeX 시스템을 거쳐 저장된다.

넷째로 구조를 가지고 있다. 각각의 문서는 `\documentstyle` 명령어로 시작하고 `\part` `\chapter` `\section` `\subsection` 과 같은 구조로 되어 있다.



예 1 문자 생성 명령의 예

Gnus of the World	<code>\wtitle{Gnus of the World}</code>
R.Dather	<code>\Wauthor{R. Dather Wand B. Falters}</code>
W.Conkright	<code>\Wauthor{Wand W.Conkright}</code>
4 July 1997	<code>\Wdata{4 July 1997}</code>

예 2 LaTeX의 문서 형태의 예

이처럼 사용자는 LaTeX의 구조에 맞게 문서를 작성해줌으로써 자신이 표현하고 싶은 것을 표현할 수 있다. 하지만 LaTeX 시스템은 고정된 환경과 많은 제약사항으로 사용하기가 쉽지 않은 단점이 있다.

5. 결론

본 논문에서는 논문 작성 시 참고 하는 참고문헌을 자동으로 삽입하기 위하여 기존의 LaTeX System에서 제공하는 참고문헌 입력 기능을 논리적인 계층 구조를 잘 나타내 줄 수 있고 구조 내용을 효율적으로 이용할 수 있는 XML을 배경으로 하여 시스템을 구현하였다. 그 결과 참고문헌 삽입작업을 위한 제한된 환경이 개선되었으며 편리한 메뉴 중심의 사용자 인터페이스는 자동 삽입 기능 외에 참고문헌 사전의 종합적인 관리와 완성된 논문의 즉각적인 웹 게시를 가능하게 되었다.

6. 참고문헌

- [1] W3C, "Extensible Markup Language 1.0," <http://www.w3.org/TR/1998/REC-XML-19980210>, 1998.
- [2] R. Davis, T. Moore, and J. Zobel, "Database Systems for Structured Document," Int'l. Symp. Advanced Database Tech and their Integration (ADTI '94), pp. 272-283, Nara Japan, 1994.
- [3] W3C, Document Object Model(Dom)Level1, Recommendation, Oct.1998.
- [4] W3C, XSL Transformations (XSLT), Recommendation, Oct. 1999. (<http://www.w3.org/TR/WD-xsl>)
- [5] Donald E. Knuth. The TEXbook. Addison-Wesley Publishing Co., Reading, Ma.,1986.