

MathML 수식 편집을 포함한 XML 문서편집 시스템의 설계 및 구현

윤화목*,⁰김철순**,정회경**
*한국과학기술정보연구원, **배재대학교 컴퓨터공학과
*hmyoon@kisti.re.kr, **cskim⁰@markup.paichai.ac.kr, hkjung@mail.pcu.ac.kr

Design and implementation of XML document edit system, inclusive of MathML mathematical formula edit

Hwa-Mook* Yoon · ⁰Cheol-Soon Kim · Hoe-Kyoung Jung**
*KISTI · **Dept. of Computer Engineering Paichai University

요 약

정보화 사회로 발전해 나아감에 따라 컴퓨터를 이용한 문서처리의 중요성이 날로 증가해가면서, 문서를 전자적으로 처리하기 위한 전자문서처리 시스템이 널리 이용되고 있는 실정이다. 하지만 기존의 문서편집 시스템들에서 수학적 표현은 이미지나 텍스트 등의 비 구조적인 방법에 의해 표현되거나 처리됨에 따라 사용자가 읽거나 표현하는데 어려움이 있다. 따라서 이러한 단점을 보완하면서 최소한의 노력으로 효과적으로 수학적 표현을 나타내는 구조적인 문서생성을 위한 노력이 필요하다. 이에 본 논문에서는 수학적 표현을 구조적으로 생성해주는 MathML(Mathematical Markup Language)의 적용이 가능한 XML(eXtensible Markup Language)기반의 구조적 문서생성을 위한 문서 편집 시스템을 설계 및 구현하였다.

1. 서론

컴퓨터를 이용한 전자문서처리에 이용되는 기존의 문서편집 시스템들은 비구조적인 문서구조를 가지고 있다. 이러한 문서를 다루는 문서편집 시스템에서 사용되는 수학적 표현은 이미지나 텍스트 등의 비구조적인 방법에 의해 표현되거나 처리된다. 이렇게 사용된 수학적 표현은 읽고 표현하기가 어려울 뿐 아니라 문서의 처리와 교환에 상대적인 불편을 초래하고 검색과 인덱싱 그리고 다른 시스템에서의 재사용성에 제한을 받게 된다.

W3C(World Wide Web Consortium)에서는 이러한 기존 시스템에서의 수식구조 표현 및 처리 등의 단점을 극복하고자 웹상에서 수학적 표현의 효율적인 표기 및 저장, 공유를 가능하게 하기 위한 MathML을 제안하였다. 그러나 올바른 MathML 표현을 위해서는 MathML의 구조 및 복잡한 MathML DTD(Document Type Definition)에 관련된 전문적인 지식을 요구하게 된다[1].

이에, 본 논문에서는 MathML의 복잡한 수식구조에 관한 전문적인 지식이 필요 없이 구조화된 문서상에 손쉽게 MathML 수식구조를 삽입할 수 있는 사용자 중심의 인터페이스를 제공하는 MathML 수식 편집기와 편집된 MathML 수식구조를 적용할 XML 문서를 편집할 수 있는 XML 문서편집 시스템을 설계 및 구현하였다[2].

본 논문의 구성은 다음과 같다. 2장에서 본 시스템의 이해와 개발에 필요한 기본 개념을 설명하고, 3장에서는 시스템의 설계 부분으로 전체 시스템의 구성을 살펴보고, XML 문서 처리부와 XML 문서 편집부, 그리고 MathML 편집부와 처리부로 나누어서 설명한다. 4장에서는 본 시스템의 구현을 설명하고, 5장에서는 결론 및 고찰, 향후 과제를 제시한다.

2. 기본 개념

2.1 XML의 개념

XML은 인터넷상에서 구조화된 문서의 교환 및 응용프로그램에서 보다 쉽게 처리하기 위해 W3C에서 1998년 표준으로 제정되었다. 이는 단순히 고정되어 있는 마크업 언어인 HTML(Hyper Text Markup Language)을 뛰어넘는 인터넷상에서 사용될 수 있는 메타 언어로서의 유연함을 보인다. MathML, CML(Chemical Markup Language) 등의 마크업 언어는 메타언어인 XML로 만들어진 좋은 예로 들 수 있다. 또한 XML은 데이터를 기술하기 위한 구조를 설계하고 이를 일반적인 문자로 작성할 수 있는 범용성을 보인다.

XML의 최종 목표는 현재 웹(Web)상에서 HTML이 처리되는 것처럼 일반적인 SGML의 처리가 가능하도록 하는 것이다.

2.2 MathML의 개념

XML 어플리케이션 중의 하나인 MathML은 W3C에 의해 1999년에 MathML 1.0이 권고안으로 발표되었고, 현재는 MathML 2.0이 2001년에 권고안으로 발표되었다. MathML은 전문적으로 수학적 표현하기 위해 설계되었고, 또한 메타언어로서 문서의 구조와 내용을 담을 수 있다. MathML의 엘리먼트는 추상적인 표기법의 구조를 기술하는 표현 마크업과 수학적 표현에 담긴 수학적 의미를 기술하는 내용 마크업으로 구성된다.

MathML의 최종목표는 수학적 표현이 웹을 통하여 표현되고 활용되는 것으로 수학적 표현과 수학적 내용을 인코딩 할 수 있는 기반기술을 제공하는 것이다.

3. 시스템 설계

본 시스템을 DTD를 이용한 Valid 문서 혹은 Well-Formed XML문서를 입력하여 편집할 수 있는 XML 문서 편집부와 새로운 XML문서 구조를 생성하거나 처리할 수 있는 XML 문서 처리부, 그리고 수학적 표현을 사용자 인터페이스를 이용해 직접 표현하면서 편집할 수 있는 MathML 문서 편집부, 마지막으로 표현된 수식구조를 XML구조 내에 삽입할 수 있는 MathML 처리부로 구성된다.

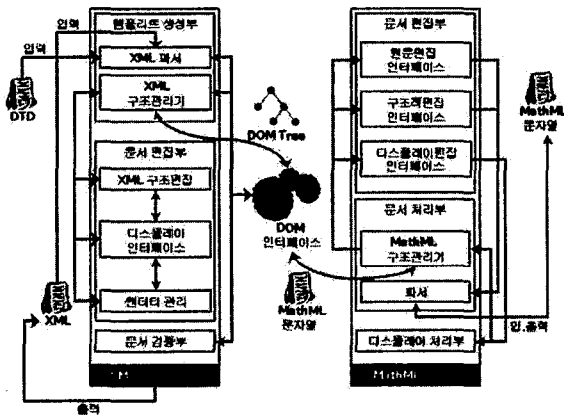


그림 1 시스템 구성도

그림 1은 본 편집시스템의 전체 구성도이다.

3.1 XML 문서 처리부

XML 문서 처리부는 XML문서를 생성하거나 기존의 문서를 새로이 편집하는 부분을 담당하게 된다. 그림 2는 XML 문서 처리부의 구성 및 처리부를 통해 관리되는 XML 처리부의 구조를 보인다.

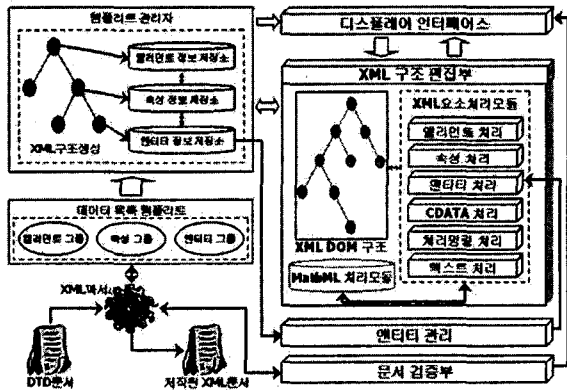


그림 2 XML 문서 처리부 구성 및 처리 구조

XML문서 처리부는 Valid문서 생성과 Well-Formed문서 생성의 두 가지 방식으로 사용된다. Valid문서 생성은 템플릿을 구성할 DTD 파일을 템플릿 생성부에 입력하게 되면 파서를 통해 DTD 파일에 대한 유효성을 검사하고 DTD에 선언된 데이터들의 목록을 작성하여 템플릿 관리자에게 전달하게 된다.

다. 템플릿 관리자에서는 전달받은 데이터들을 이용해 XML 구조를 생성하게 되고, 이렇게 생성된 XML의 구조는 DOM(Document Object Model) 객체 트리로 구성되고 이것을 이용하여 디스플레이 관리자는 트리 구조 인터페이스와 소스 편집 인터페이스를 이용하여 XML문서구조를 편집할 수 있는 정보를 만들어준다[3].

Well-Formed 문서 생성은 저장된 XML문서를 템플릿 생성부에 입력하게 되면 파서를 통해 유효성을 검사하고 템플릿 목록을 작성하여 템플릿 관리자에게 넘겨주며 나머지 처리는 Valid 문서 생성의 경우와 같이 처리된다.

3.2 XML 문서 편집부

문서 편집부는 XML의 텍스트 노드에 대해 편집기능을 제공하는 CRichEditCtrl 클래스를 상속받는 디스플레이 편집 인터페이스와 논리적인 계층구조를 표현하고 이에 대한 편집기능을 제공하는 CTreeCtrl 클래스를 상속받는 XML 구조편집 인터페이스, 엔티티 정보들이 디스플레이 되어 XML 노드 사이에 직접 삽입이 가능한 기능을 제공하는 CListCtrl 클래스에 상속된 엔티티관리 인터페이스로 구성된다. 각 인터페이스를 통해 편집되는 데이터는 문서 처리부로부터 관리되고 정의되는 XML 구조를 통해 처리되며, 각 인터페이스의 특성에 맞는 데이터와 XML 구조간의 처리를 통해 갱신이 이루어 지고, 구조 편집 인터페이스와 디스플레이 편집 인터페이스의 데이터들 사이에 동기화가 이루어진다.

3.3 MathML 문서 편집부

MathML 문서 편집시스템은 임출력과 편집을 통해 처리되고 변경되는 여러 데이터들을 처리, 관리하고 수학적 디스플레이 처리를 담당하고 사용자 인터페이스의 제공을 담당하는 문서 편집부와 처리된 수학적 구조를 XML 문서 편집 시스템으로 전달하는 MathML 처리부로 크게 나뉜다. 그림 3은 MathML 문서 편집부의 구성 및 처리 구조를 보인다.

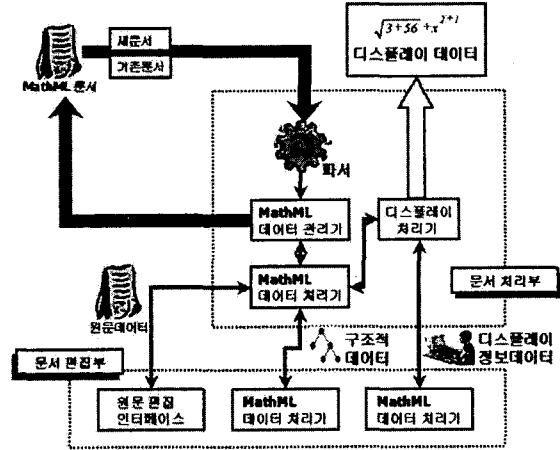


그림 3 MathML 문서편집부의 구성 및 처리구조

로드되거나 초기화된 MathML 문서는 파서를 통해 W3C 표준 인터페이스인 DOM으로 구성되며 MathML 데이터 관리기를 통해 관리된다. DOM객체는 MathML 데이터 처리기에 의해 변환 과정을 거쳐, 입력된 MathML문서의 논리적인 계층구조를 표현하는 트리형태로 표현해주고, MathML문서가 의미하

는 수학적 표현을 디스플레이 처리기를 통해 디스플레이 해주게 되면, 사용자는 이러한 정보들과 시스템이 제공하는 편집 인터페이스를 통해 원하는 형태의 편집을 할 수 있게 된다.

3.4 MathML 처리부

본 시스템에서 처리되는 모든 수학적식은 MathML 문서 편집기를 통해서 처리된다. 그림 4는 MathML 데이터 처리부의 구성도 및 처리과정을 보인다.

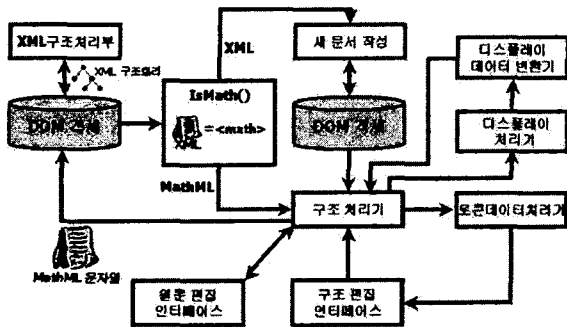


그림 4 MathML 처리부의 구성 및 처리과정

XML DOM 으로부터 받은 XML 노드를 IsMath() 함수를 이용하여 MathML 태그이면 구조 처리기로 MathML 구조를 전달해주고 MathML 태그가 아니면 새 문서 작성으로 판단하여 DOM에 전달해준다. 구조 처리기를 통한 MathML의 정보는 사용자 편집에 의해 다시 새로운 구조로 변환되고 이 결과를 XML 문서편집기의 DOM 객체에 문자열로 전달한다. 수식구조를 전달 받은 XML문서 편집기는 DOM 객체를 이용해 새로운 XML구조로 생성하여 기존 노드에 삽입하게 된다.

4. 구현

4.1 XML 문서편집 시스템 및 MathML 편집 시스템

본 시스템은 Windows 2000 Server 환경에서 Visual C++ 6.0으로 개발하였다. 파서는 Microsoft사의 MSXML4.0을 이용하였다.

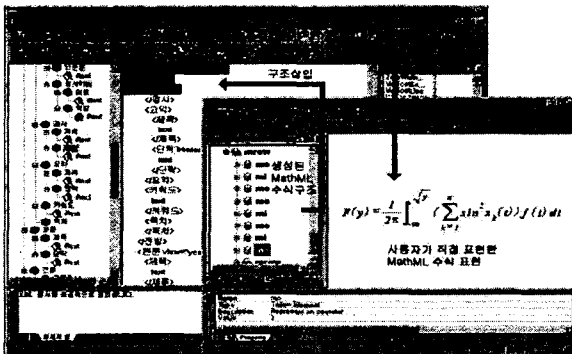


그림 5 XML 문서편집 시스템 및 MathML 편집 시스템

그림 5는 XML 문서 편집 인터페이스와 MathML 편집 인터페이스의 구현 화면이다. 그림 5의 우측 상단에 있는 MathML 삽입 인터페이스를 이용해 MathML 편집기를 호출하면 사용자가 직접 수식을 표현함에 따라 새로운 MathML 수식구조를 생성하게 되고 MathML 편집기를 종료함으로써 생성된 수식구조를 그림 6과 같이 XML 문서구조 내에 삽입하게 된다.

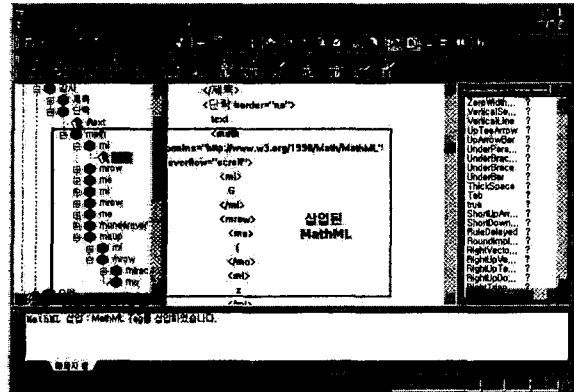


그림 6 MathML 삽입

5. 결론 및 고찰

본 논문에서는 문법에 맞게 작성된 기존의 MathML의 구조는 곧바로 XML문서에 적용시킬 수 있으며, MathML의 복잡한 구조를 알지 못하더라도 수학적 표현을 쉽게 MathML의 구조로 작성하여 XML문서에 적용시킬 수 있는 효율적인 인터페이스를 제공하였다. 뿐만 아니라 기존 XML편집기의 일반적인 문서 편집 방식의 단점인 XML 구조 소스를 일반 사용자가 직접 편집하게 됨으로써 발생할 수 있는 오류를 최소화 하기 위한 대안으로 템플릿 관리자를 제공함으로써 문서의 무결성을 유지시켰다.

본 편집 시스템은 XML을 기반으로 하고 수학적 표현해아 하는 e-book이나 수많은 관련 어플리케이션의 효율적인 처리에 유용하게 사용될 것이며, 웹 상에서 MathML 사용의 보편화에 기여하게 될 것이다.

본 시스템에서 연구되어야 할 기능적인 측면에서는 외부 파일에 접근하게 해주는 링크기능이나 PUBLIC 처리 등의 기능이 추가 되어야 하며, 표현적인 측면에서는 그림 엔티티와 테이블의 표현이나 XML편집기 상에서 수학적 직접적인 표현 등의 연구가 향후 연구과제로 남는다.

6. 참고 문헌

- [1] W3C's Math Home Page, <http://www.w3.org/Math/>
- [2] W3C, Extensible Markup Language(XML) Version 1.0, <http://www.w3.org/TR/REC-xml>, Feb. 10, 1998
- [3] W3C's Document Object Model, <http://www.w3.org/DOM>
- [4] Amaya - W3C's Editor/Browser, <http://www.w3.org/Amaya/>