

---

# MathML 수식 구조 표현을 지향하는 XML 문서 편집 시스템의 설계 및 구현

김철순\* · 정회경\*

\*배재대학교 컴퓨터공학과

Design and implementation of XML document edit system that intend to  
MathML mathematical formula structure representation

Cheol-Soon Kim · Hoe-Kyoung Jung

\*Dept. of Computer Engineering Paichai University

E-mail : \*cskim@markup.paichai.ac.kr, \*hkjung@mail.pcu.ac.kr

## 요 약

컴퓨터를 이용한 전자문서처리에 이용되는 기존의 문서편집 시스템에서는 비구조적인 문서구조를 다루게 됨으로써 시스템 내에서 사용되는 수학적 표현은 이미지나 텍스트 등의 비구조적인 방법에 의해 표현되거나 처리된다. 이렇게 사용된 수학적 표현은 문서의 가독성과 재사용성 그리고 문서의 처리와 교환에 상대적인 불편을 초래하게 된다. 그러므로 이러한 단점을 극복하고 효율적으로 구조적인 문서상에 MathML(Mathematical Markup Language) 수식 구조를 적용시킬 수 있는 문서편집 시스템이 요구된다. 이에 본 논문에서는 MathML을 기반으로 한 수학적 표현을 구조적으로 처리할 수 있는 MathML 수식 편집이 가능한 XML(eXtensible Markup Language) 기반의 구조적 문서생성을 위한 문서 편집 시스템을 설계 및 구현하였다.

## ABSTRACT

Represent of mathematical formula used within system handling document that is nonstructural in existent document editing system that is used in electron document processing that use computer is represented or processed by method that is nonstructural of image or text or etc. Such mathematical formular causes relative inconvenience to readability and reusability of document and processing and exchange of document. Therefore, document editing system is required that can overcome such nonadvantage and apply MathML mathematical formula structure on efficiently structural document. Therefore, designed and implemented that document editing system for structural document creation of XML base that can mathematical formular editing of MathML base in this paper.

## 키워드

MathML, XML, 수식편집기, XML편집기, math, mathematical

## 1. 서 론

컴퓨터 관련 분야의 기술진보는 문서의 전자화를 가속화 시켰다. 그러나 이러한 전자문서 처리를 위한 기존의 시스템들은 각기 독자적인 문서구조와 표현정보를 포함하여 시스템들 간의 문서 공유에 상당한 어려움이 발생한다. 그러므로, 이러한 전자문서를 처리

하는 시스템들에서 표현되는 수학적 표현은 이미지나 텍스트 등의 비 구조적인 방법에 의해 표현되거나 처리됨에 따라 사용자들이 읽거나 표현하는데 많은 어려움이 따르게 되고, 이미지 특성상 상대적으로 큰 데이터 용량으로 인해 처리와 교환에 부담을 주게 되며, 검색이나 인덱싱에 있어서 많은 제한을 받게 된다. 이러한 문제점을 보완하기 위하여 W3C(World Wide Web

Consortium)에서는 이러한 기존 시스템에서의 수식구조 표현 및 처리 등의 단점을 극복하고자 웹 상에서 수학적 표현의 효율적인 표기 및 저장, 공유를 가능하게 하기 위한 MathML을 제정하였다. 그러나 제대로 MathML을 사용하기 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 손쉽게 MathML 수식구조를 생성할 수 있는 사용자 중심의 인터페이스를 제공하는 MathML 수식 편집기와 편집된 MathML 수식구조를 구조화된 문서상에 적용시킬 XML 문서를 편집할 수 있는 XML 문서편집 시스템을 설계 및 구현하였다[2].

본 논문은 다음과 같이 구성된다. 본문의 1장에서 본 시스템의 이해와 개발에 필요한 기본 개념을 설명하고, 2장에서는 시스템의 설계 부분으로서 전체 시스템의 구성을 살펴보고, 전체 시스템을 XML 문서 입력 및 생성모듈과 XML 문서 편집 및 처리모듈, 그리고 MathML 구조 생성모듈과 MathML 편집모듈 그리고 MathML 삽입모듈로 나누어서 설명한다. 3장에서는 본 시스템의 구현을 설명하고, 마지막으로 결론 및 고찰, 향후 연구 과제를 제시한다.

## II. 본 론

### 1. MathML의 기본 개념

XML 어플리케이션 중의 하나인 MathML은 W3C에 의해 1999년에 MathML 1.0이 권고안으로 발표되었고, 현재는 MathML 2.0이 2001년에 권고안으로 발표되었다. MathML은 전문적으로 수학적 표현을 위해 설계되었고, 또한 메타언어로써 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 식이 웹을 통하여 표현되고 활용되는 것으로 수학적 표현과 수학적 내용을 인코딩 할 수 있는 기반기술을 제공하는 것이다.

### 2 시스템 설계

본 시스템은 DTD문서 혹은 XML문서를 입력하여 XML문서의 구조를 생성할 수 있는 XML 문서 입력 및 생성모듈과 XML문서 구조를 변환하거나 XML 노드들을 처리할 수 있는 XML 문서 편집 및 처리모듈, 그리고 MathML 문서를 생성하거나 입력받을 수 있는

MathML 생성모듈과 사용자 중심의 인터페이스를 이용해 MathML 수학적 식의 직접 표현이 가능하고 표현된 MathML 수학적 식을 편집할 수 있는 MathML 문서 편집 모듈, 마지막으로 표현된 수식구조를 XML구조 내에 삽입할 수 있는 MathML 삽입모듈로 구성된다.

그림 1은 본 편집시스템의 전체 구성도이다.

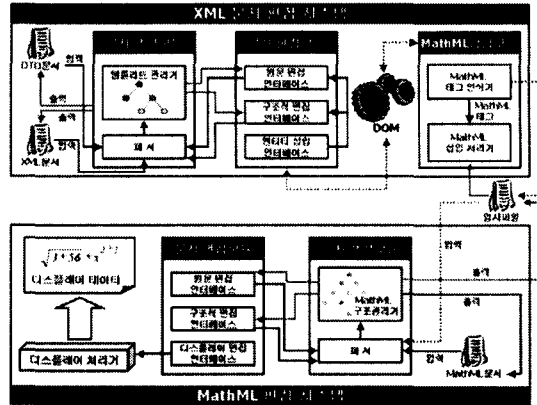


그림 1 전체 시스템 구성도

### 2.1 XML 문서 입력 및 생성모듈

XML 문서 입력 및 생성모듈은 XML문서를 생성하거나 기존의 문서를 입력받아 새로이 XML 템플리트를 구성하는 부분을 담당하게 된다. 그림 2는 XML 문서 입력 및 템플리트 관리기를 통해 관리되는 XML 생성모듈의 구조를 보인다.

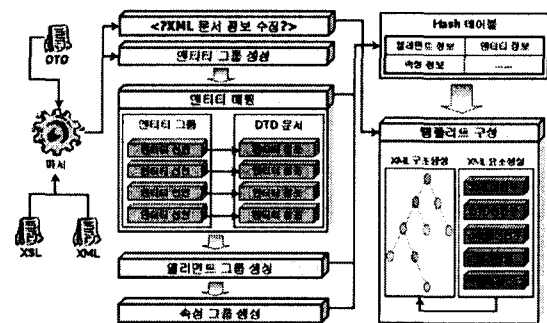


그림 2 XML문서 생성모듈의 구성 및 처리

문서 생성모듈은 크게 Well-Formed XML 문서생성과 Valid XML 문서생성의 두 가지 방식으로 사용된다. Valid 문서생성은 템플리트를 구성할 DTD 파일을 XML 템플리트 생성자에 입력하면 파서를 통해 DTD 파일에 대한 정형성을 검사하고 DTD에 선언된 XML노드 데이터들의 목록을 작성하여 템플리트 관리자에게 전달하게 된다. 템플리트 관리자에서는 전달받은 데이터들을 이용해 새문서 작성의 경우에는 새로운 XML구조를 생성하게 되고, 기존 문서가 입력된 경우

에는 기존 문서와 DTD 파일과의 유효성 검사를 실시한 후, 입력된 XML 문서의 구조를 메모리에 생성시키게 된다. 이렇게 생성된 XML의 구조는 DOM(Document Object Model) 객체 트리로 구성되고 이것을 이용하여 인터페이스 관리자는 트리 구조 인터페이스와 소스 편집 인터페이스를 이용하여 XML 문서구조를 편집할 수 있는 정보를 만들어준다[3].

Well-Formed XML 문서생성은 기존의 저작된 XML문서의 경우에 XML 구조를 템플릿 생성자에 입력하게 되면 파서를 통해 문서의 정형성을 검사하고 템플릿 목록을 작성하여 템플릿 관리자에게 넘겨주며 새로운 XML문서를 작성할 때는 템플릿 생성자에서 기본 XML구조(XML선언부, 최상위노드)만을 생성한 후 템플릿 관리자에게 넘겨주게 되며 나머지 처리는 Valid XML 문서 생성의 경우와 동일하게 처리된다.

### 2.2 XML 문서 편집 및 처리모듈

문서 편집 및 처리모듈은 XML의 PCDATA 노드에 대해서만 편집기능을 제공하는 CRichEditCtrl클래스를 상속받는 CTextEditView 클래스인 XML 원문 편집 인터페이스와 XML의 논리적인 계층구조를 표현하고 이에 대한 구조편집 기능을 제공하는 CTreeCtrl 클래스를 상속받는 CNodeTreeView 클래스인 XML 구조편집 인터페이스, 엔티티 정보들이 디스플레이 되어 XML 노드사이에 직접 삽입이 가능한 기능을 제공하는 CListCtrl 클래스에 상속된 엔티티 삽입 인터페이스로 구성된다. 각 인터페이스를 통해 편집되는 데이터는 XML 구조편집 클래스인 CDwDOMCtrl 클래스로부터 관리되고 정의되는 XML 구조를 통해 처리되며, 각 인터페이스의 특성에 맞는 데이터와 XML 구조간의 처리를 통해 문서구조의 갱신이 이루어지고, 구조 편집 인터페이스와 디스플레이 편집 인터페이스의 데이터들 사이에 동기화가 이루어진다.

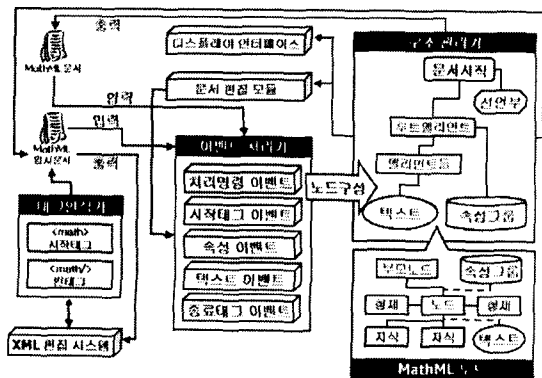


그림 3 MathML 문서 생성 모듈의 구성 및 처리

### 2.3 MathML 문서 생성모듈

그림 3은 MathML 문서 생성모듈의 구성 및 처리 구조를 보인다.

MathML 문서 편집 시스템은 파일 입력과 문서 구조를 생성하는 문서 생성 모듈과 편집을 통해 처리되고 변경되는 여러 데이터들을 처리, 관리하는 문서 편집모듈 그리고 처리된 수학적 구조를 XML 문서 편집 시스템으로 전달하는 MathML 삽입 모듈로 크게 나뉜다.

그림 3과 같이 로드되거나 초기화된 MathML 문서는 파서를 통해 W3C 표준 인터페이스인 DOM 객체로 구성되며 MathML 데이터 관리기를 통해 관리된다. DOM객체는 MathML 데이터 처리기에 의해 변환 과정을 거쳐, 입력된 MathML 문서의 논리적인 계층 구조를 표현하는 트리형태로 표현해주고, MathML 문서가 의미하는 수학적 표현을 디스플레이 처리기를 통해 디스플레이 해주게 되면, 사용자는 이러한 정보들과 시스템이 제공하는 편집 인터페이스를 통해 원하는 형태의 편집을 할 수 있게 된다.

### 2.4 MathML 편집모듈

문서 편집부는 MathML의 기본적인 원문 텍스트의 편집기능을 제공하는 원문 편집 인터페이스와 논리적인 계층구조를 표현하고 이에 대한 편집기능을 제공하는 구조적 편집 인터페이스, 수학적식이 디스플레이 되는 부분을 직접 편집하는 기능을 제공하는 디스플레이 편집 인터페이스로 구성된다. 각 인터페이스를 통해 편집되는 데이터는 문서 편집모듈로부터 관리되고 정의되는 MathML 노드 구조를 통해 처리되며, 각 인터페이스의 특성에 맞는 데이터와 MathML 노드구조간의 변환을 통해 갱신이 이루어지고, 각 편집 인터페이스 데이터간의 동기화가 이루어진다.

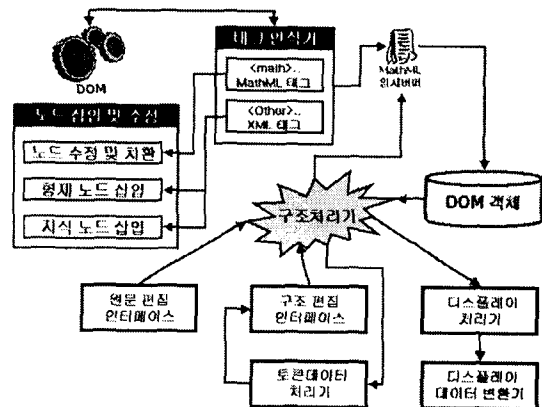


그림 4 MathML 삽입 모듈의 구성 및 처리

### 2.5 MathML 삽입모듈

본 시스템에서 처리되는 모든 수학식은 MathML 문서 편집 및 처리모듈을 통해서 처리된다. 그림 4는 MathML 삽입모듈의 구성도 및 처리과정을 보인다.

DOM 객체로부터 전달받은 XML 노드를 XML 편집기의 IsMath() 함수를 이용하여 MathML 태그의 존재 여부를 판단하여 MathML 태그가 존재하면 외부 임시 파일로 저장한 후, MathML 편집기가 실행될 때 저장된 파일을 열어 그 내용을 구조 처리기로 전달하여 새로운 구조로 편집할 수 있는 기본 정보를 생성해준다. 만약, MathML 태그가 존재하지 않으면 새로운 문서 작성으로 판단하여 MathML 편집기만을 실행하게 된다. 구조 처리기에 의해 처리된 MathML 구조는 사용자 편집에 따라서 다시 새로운 구조로 변환되고, 이 결과를 XML 문서 편집기의 DOM 객체에 문자열로 전달한다. 문자열을 전달받은 XML 문서 편집기는 DOM 객체를 이용해 새로운 XML 노드로 생성되고 이 노드를 이용하여 사용자가 지정한 위치에 해당 노드를 삽입하게 된다.

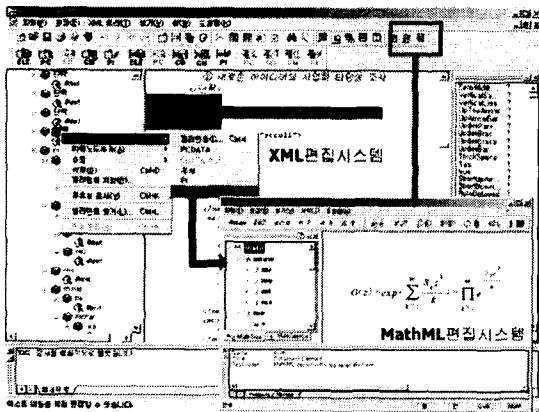


그림 5 XML과 MathML 편집시스템의 구현화면

### 3. 구현

본 시스템은 Windows 2000 Server 환경에서 Microsoft Visual C++ 6.0으로 개발하였고, XML과 MathML의 데이터 처리 및 관리를 위한 파서는 Microsoft MSXML4.0을 사용하였다. 템플릿 관리자는 DLL(Dynamic Linking Language)로 개발하여 다른 시스템에서도 사용이 가능하도록 구현하였다.

그림 5는 XML 문서 편집 시스템과 MathML 문서 편집 시스템의 구현 화면이다.

XML 편집 시스템의 우측 상단에 있는 MathML 삽입 인터페이스를 이용해 MathML 편집 시스템을 호출하면 MathML 편집 시스템의 디스플레이 인터페이스를 이용하여 사용자가 직접 수식을 편집함에 따라 표

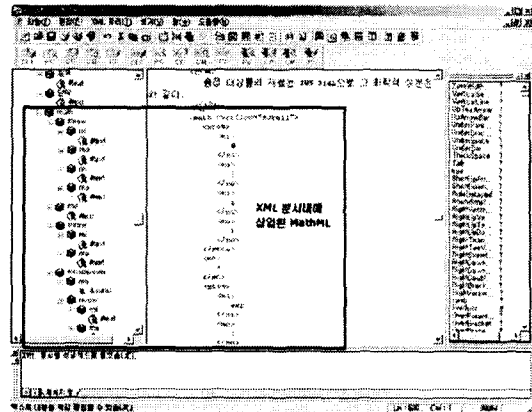


그림 6 MathML 삽입 결과화면

현된 수식에 해당하는 새로운 MathML 수식구조를 생성하게 되고 MathML 편집 시스템을 종료함으로써 생성된 수식구조를 XML 편집 시스템에서 XML 문서구조의 특정위치 내에 삽입하게 된다.

그림 6은 XML 문서구조 상에 MathML 태그가 삽입된 결과를 보여준다.

### III. 결론

본 논문에서는 MathML의 기본 문법과 복잡한 구조정보에 대한 기본 지식이 없어도 수학식의 표현을 쉽게 MathML의 구조로 작성할 수 있는 효과적인 방법을 제안하였고, 문법에 맞게 작성된 기존의 MathML의 구조는 곧바로 XML 문서에 적용시킬 수 있는 효율적인 인터페이스를 제공하였다. 뿐만 아니라 기존 XML 편집기의 일반적인 문서 편집 방식의 단점인 XML 구조 소스를 일반 사용자가 직접 편집하게 됨으로써 발생할 수 있는 오류를 최소화하기 위한 대안으로 템플릿 관리자를 제공함으로써 문서의 무결성을 유지시켰다.

본 편집 시스템은 XML을 기반으로 하고 수학식을 표현해야 하는 e-book이나 수많은 관련 어플리케이션의 효율적인 처리에 유용하게 사용될 수 있을 뿐만 아니라, 웹 상에서 MathML 사용의 보편화에 기여하게 될 것이다. 향후 연구되어야 할 기능적인 측면으로는 외부 파일에 접근하게 해주는 링크기능이나 URI(Uniform Resource Identifiers) 처리기능과 XML 문서의 문자열 인코딩 처리와 디코딩기능 등이 추가되어야 하며, 표현적인 측면에서는 그림 엔티티와 테이블의 표현이나 XML 편집기 상에서 수학식의 직접적인 표현 등의 연구가 필요하다.

### 참고문헌

- [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>