

GUI를 기반으로 하는 SGML DTD

생성기의 설계 및 구현

한상훈^o, 이만형, 박세원, 황대훈
경원대학교 전자계산학과

Design and Implementation of SGML DTD Generator based on GUI

SangHoon Han^o, ManHyung Lee, SeWon Park, DaeHoon Hwang

Dept. of Computer Science, Kyungwon Univ.

요약

이종간 간에 문서 정보를 상호 교환할 수 있도록 국제표준기구에서는 SGML을 데이터 객체 양식 표준으로 ISO 8879로 제정하였다. 오늘 날 웹 표준 문서인 HTML이 XML로 전환되어 감에 따라 SGML의 중요성이 증대되고 있다. HTML이나 XML은 일반사용자가 사용하기 쉽게 SGML의 복잡함을 축소한 것이다. XML이나 SGML은 모두 DTD 규칙에 의해 작성되는데, DTD를 작성하기 위해서는 SGML DTD 작성 규칙을 알아야 한다. 이에 본 논문에서는 SGML DTD 작성 규칙을 잘 모르더라도 쉽게 DTD를 개발할 수 있도록 GUI를 기반으로 하는 DTD 생성기를 구현하였다.

I. 서론

정보화 사회에서 정보의 개방 환경에서 컴퓨터를 이용한 문서 처리와 문서를 이 기종간의 시스템에서 상호 교환하는 정보 교환의 중요성이 날로 증대되고 있다. 이에 따라 텍스트, 그래픽스, 오디오 및 비디오 등을 포함하는 멀티미디어 문서들을 효율적으로 교환 및 저장하고, 응용, 주변장치, 네트워크로부터 독립적으로 문서를 처리할 수 있는 환경이 요구되고 있다.

이에 이 기종간의 문서 교환 및 처리를 목적으로 문서의 논리적 구조를 효율적으로 제공하는 SGML(Standard Generalized Markup Language)이 문서의 마크업 언어 구성을 위한 표준안으로 국제 표준화 기구(International Standard Organization)에서 제정되었다[3]. 현재 HTML(HyperText Markup Language), HyTime(Hypermedia Time based Structuring Language), TEI(Text Encoding Initiative)등이 SGML을 이용하여 구성되었으며, CALS 규격 내에서도 SGML을 문서 처

리를 위한 표준안으로 제정하고 있다. 또한 현재 널리 연구 개발되고 있는 인트라넷에서도 SGML의 처리 기술이 절실히 요구되고 있는 실정이다[9].

이에 따라 본 논문에서는 SGML에서 규정하는 DTD(Document Type Definition) 작성 규칙에 따라 DTD를 생성 및 편집할 수 있고, 한글 처리가 가능한 DTD 생성기를 설계 및 구현하였다.

DTD는 일반 텍스트 에디터로 작성이 가능하지만 문서의 논리적 구조를 파악하기 어렵기 때문에 본 논문에서는 문서의 논리적 구조를 파악하기 쉽게 문서의 논리적 구조를 트리 형식으로 설계하였고, 사용자가 DTD의 작성 규칙을 자세히 모르더라도 GUI(Graphic User Interface)를 기반으로 하여 쉽게 DTD 트리를 작성하는 것으로 DTD를 작성할 수 있도록 구현하였다.

DTD를 작성한 뒤 다른 SGML 응용 프로그램에서 사용이 가능하도록 SGML 형식으로 DTD 파일을 생성하도록 하였으며, 이미 생성된 DTD 파일을 불러서 편집이 가능하며, 생성된 DTD 파일을 사용자가 텍스트

에디터 형식으로 편집할 수 있도록 텍스트 편집 기능을 두었다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구에 대해 알아보고 3장에서는 본 연구에서 설계 및 구현한 SGML DTD 생성기의 구성에 대해 논의한다. 끝으로 4장에서는 결론에 대해 정리하였다.

II. 관련연구

2.1 SGML

SGML은 텍스트, 그래픽, 오디오 및 비디오 등을 포함하는 전자 문서들을 이 기종의 시스템간에 효율적으로 전송, 저장 및 자동 처리하기 위한 ISO 권고 표준이다. SGML은 문서의 논리적인 구조를 기술하기 위한 목적으로 구성되었으며 문서의 레이아웃을 기술하는 표준안인 DSSSL과 결합하여 문서를 효과적으로 표현할 수 있도록 지원한다.

SGML 문서는 해당 업무에 따라 문서의 구조를 정의한 DTD를 통하여 문서의 구성 요소간의 관계와 규칙을 정한다[1].

이와 같이 SGML은 그 자체가 마크업 언어라기 보다는 마크업 언어를 개발하는 데 사용되는 언어로서, SGML로 기술된 대표적인 마크업 언어가 HTML, HyTime 등이다. SGML 문서는 마크업 방식을 선언하는 부분과 실제 문서가 결합된 형태로 존재하는데, 경우에 따라서는 문서의 레이아웃 정보를 가진 DSSSL 정보가 결합되어 함께 제공되기도 한다. SGML 문서는 크게 나뉘어 SGML 선언부, SGML 문서형 정의부, SGML 실제문서부로 구성되어 있다[2][4].

2.2 SGML 문서 처리 과정

일반적으로 많이 사용하는 워드 프로세서 및 DTP 환경에서는 문서 내에 프리젠테이션 정보가 삽입되어 있으며 문서의 구조 정보는 포함되지 않는다. 또한 워드 프로세서마다 독자적인 프리젠테이션 포맷을 사용한다. 따라서 이러한 문서는 이 기종간의 정보의 공유 및 교환을 어렵게 하며 문서 내에 프리젠테이션 정보가 결합되어 있기 때문에 재사용 측면에서 여러 가지 문제점을 가지고 있다. 이러한 이유로 인하여 SGML에서는 문서의 내용과 구조 정보로부터 프리젠테이션 정보를 분리하여 문서를 특정 소프트웨어나 하드웨어에 독립적이고 다양한 용도에 맞게 재 사용할 수 있도록 한다. 그림 1은 SGML 문서의 처리과정을 보여주고 있다.

SGML 문서에는 문서의 레이아웃이나 프리젠테이션에 관한 정보를 포함하지 않기 때문에 SGML 문서를 프리젠테이션하기 위해서는 이에 대한 정보를 별도로 제공하여야 한다. 이를 위해 ISO에서는 SGML 문서의

다양한 처리정보를 기술하기 위한 표준으로 DSSSL (Document Style Semantics and Specification 언어 : ISO 10179)을 정의하였다.

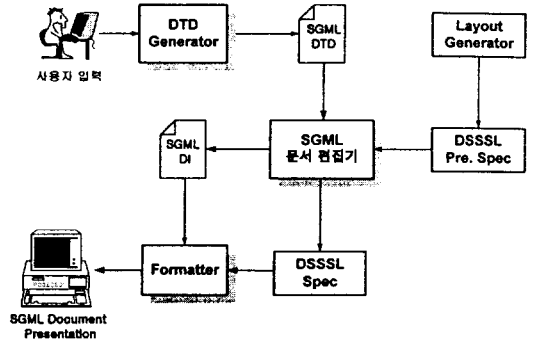


그림 1 SGML 문서 처리 과정

SGML 문서 처리 과정은 사용자가 문서의 논리적 구조 정보를 가지고 있는 DTD의 개발이 필요하다. DTD 생성기나 텍스트 에디터로 DTD를 개발하고, SGML 편집기로 SGML 실제 문서부를 작성한다. 이때 SGML 실제 문서부는 개발한 DTD 규칙대로 작성하여야 한다. SGML 실제 문서부와 DTD는 문서의 프리젠테이션에 관한 정보를 포함하지 않으므로, 문서의 프리젠테이션 정보를 포함하는 DSSSL 정보와 결합하여 프리젠테이션하게 된다.

2.3 SGML DTD

SGML 실제 문서 부를 만들기 위해서 문서의 구조를 정의하는 DTD는 주요 구성요소인 엘리먼트(element), 어트리뷰트(attribute), 엔티티(entity)로 작성된다[8].

DTD는 문서의 논리 구조와 여기에 관련된 속성이 선언되며, 자주 사용하는 문자열을 참조하기 위한 엔티티 선언, 주석과 표기법을 선언할 수 있다.

2.3.1 엘리먼트 선언

엘리먼트 선언(element declaration)은 SGML 문서의 논리 구조의 구성 요소를 선언하는 것으로서 식별자와 내용부로 이루어지며 부 엘리먼트를 포함할 수 있다. 부 엘리먼트를 선언하면 부 엘리먼트에 대한 선언 또한 필요하며, 트리 형태의 구조를 갖는 문서형 정의부에 있는 엘리먼트 선언들에 의해 문서의 구조를 파악하게 된다.

2.3.2 엔티티 선언

엔티티 선언(entity declaration)은 특수문자나 긴 문자열을 간단한 표기법으로 대체할 때 또는 기계에 특정

한 프로세싱을 하도록 명령할 때 선언 하는 것이다. 엔티티 선언은 컴퓨터 시스템에서 특정 파일의 포인터로 쓸 때 매우 유용하다.

2.3.3 어트리뷰트 선언

어트리뷰트 선언(attribute declaration)은 특정 엘리먼트에 대한 추가적인 설명 정보나 값을 지정해야 할 때 사용하는 선언이다. 속성은 한 엘리먼트를 다른 엘리먼트나 다른 엘리먼트 설명 데이터에 결합하는 상호 참조 정보를 포함하여, 문서의 외부 개체와 관련한 정보를 포함하는데 유용하다.

III. DTD 생성기의 설계 및 구현

3.1 시스템 구성

본 논문에서 설계 및 구현한 SGML DTD 생성기는 한글 태그를 지원할 뿐만 아니라 GUI를 기반으로 하여 일반 사용자도 쉽게 SGML DTD를 생성 및 편집할 수 있도록 인터페이스를 문서의 전체 구조를 파악하기 쉽게 트리 형태로 설계 및 구현하였다.

본 논문의 SGML DTD 생성기는 그림 2에서와 같이 크게 DTD 해석 모듈, DTD Tree 생성 모듈, DTD Tree 편집 모듈, DTD 생성 모듈로 구성되어 있다.

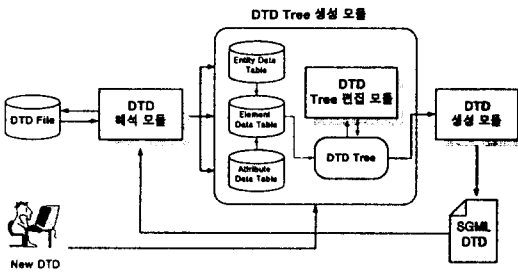


그림 2 SGML DTD 생성기의 시스템 구성도

본 논문에서 개발한 SGML DTD 생성기는 다음과 같은 순서로 수행된다.

이미 DTD 파일이 존재하는 경우에는 DTD 해석 모듈을 거쳐 이미 만든 DTD가 유효한 지를 검증하여 사용자에게 오류가 있으면 오류가 있는 부분을 출력하여 준다. DTD가 유효하면 DTD 트리 생성 모듈에서 엔티티 데이터 테이블(Entity Data Table), 엘리먼트 데이터 테이블(Element Data Table)과 어트리뷰트 데이터 테이블(Attribute Data Table)을 만들어 DTD 트리를 생성하여 사용자에게 보여주면 DTD 트리를 사용자가 편집하게 된다.

DTD 트리가 생성되면 DTD 생성 모듈은 DTD 트리를 SGML DTD 규격에 맞게 DTD를 생성하여 준다.

생성된 DTD는 다시 한번 DTD 해석 모듈을 거쳐서 DTD가 유효한지 검사하여 SGML 문법에 맞게 작성되었으면 *.dtd 파일로 최종 출력되게 된다.

DTD를 새로 생성하는 경우에는 DTD 해석 모듈을 거치지 않고 바로 DTD 트리 생성 모듈에서 DTD 트리를 작성하게 된다. 사용자가 DTD 트리 생성을 완료하면 이미 DTD 파일이 존재하는 경우와 같은 흐름을 가진다.

3.2 DTD 해석 모듈

DTD 해석 모듈은 생성 또는 편집을 통하여 작성된 DTD를 엘리먼트, 엔티티, 어트리뷰트 별로 분류하여, DTD가 SGML 문법에 맞게 작성되었는지의 오류를 체크하고, 엘리먼트 간의 연결관계를 확인하는 작업을 하는 모듈이다. 그림 3은 DTD 해석 모듈의 자료 흐름도이다.

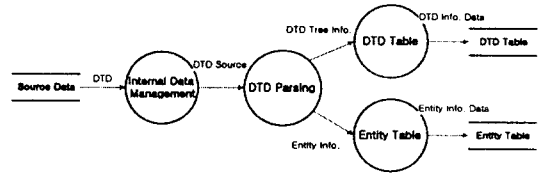


그림 3 DTD 해석 모듈의 자료 흐름도

내부 데이터 관리 모듈의 기능은 입력된 DTD 데이터를 관리해주는 기능을 하고 DTD 파싱 모듈의 기능은 입력된 DTD 데이터에서 DTD 파싱 정보를 생성하고, DTD 문법 체크를 하고 파싱 정보를 출력한다.

DTD 테이블 모듈은 DTD의 엘리먼트, 엔티티, 어트리뷰트를 분리하여 데이터 테이블을 만들어서 DTD 트리 생성 모듈에게 전달한다.

3.3 DTD 트리 생성 모듈

본 논문에서는 DTD의 문서구조를 사용자가 보기 쉽게 하기 위하여 DTD의 문서구조를 트리 형식으로 설계하였다. DTD 트리 생성 모듈은 DTD의 구성요소인 엘리먼트들을 사용자에게 보여주기 위한 기능을 하는 모듈이다. DTD 트리 생성 모듈은 MS-VISUAL C++ 라이브러리인 MFC의 CTreeView Class를 계승받아 프로그래밍 하였다. DTD 트리는 각 엘리먼트들을 보여 주게 되는데 각 트리 노드들을 연결할 때 DTD의 연결자를 표시하도록 하였으며 또 각 트리의 상위 노드와 하위 노드를 연결하기 위해 연결자가 표시되는 위치에서 노드를 연결하였다. 그림 4는 DTD 트리의 표현을 나타낸 것이다

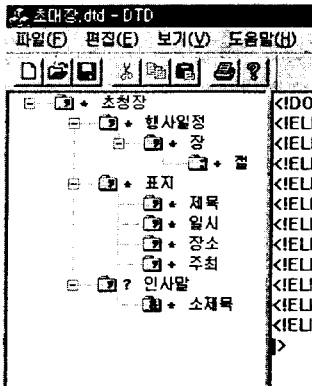


그림 4 DTD 트리의 표현

(1) 연결자의 처리

DTD의 구조는 각 엘리먼트의 관계로 보여지게 되는데 이러한 엘리먼트 들은 하나의 연결자로서 연결하게 된다. 연결자는 ‘;’, ‘&’, ‘|’로 구성되어 있는데, 각 연결자는 다음과 같은 특성을 가진다.

- , : 순서 (“~ 다음에”) 라는 의미
- & : 그리고 (모든 항목에 나타나야 하지만 순서와 상관없다)
- | : 또는 (연결자 전후에 기술된 부 엘리먼트 중 하나만 발생해도 된다)

표 1은 본 논문에서 연결자의 처리를 화면에 어떻게 표현하였는지를 나타낸다.

표 1 연결자의 화면 표현방법

연결자의 종류	표현 방법
,	
&	

(2) 사건지시자의 처리

사건 지시자는 엘리먼트의 발생여부를 표현해 주는 표시자로서 ‘?’ , ‘+’ , ‘*’ 로 구성되며, 각 사건 지시자는 다음과 같은 특성을 가진다.

- ? : 옵션 (부 엘리먼트가 나타날 수 있고 없을 수도 있다)
- + : 필수이며 반복적 (한 번 이상 발생해야 하며 반복해서 발생할 수 있다)
- * : 옵션이며 반복적 (0 번 이상 발생할 수 있다)

표 2는 사건지시자의 처리를 화면에 어떻게 표현하였는지를 나타낸다.

표 2 사건지시자의 화면 표현 방법

사건 지시자의 종류	표현 방법
?	
+	
*	

3.4 DTD Tree 편집 모듈

편집 모듈은 DTD 트리를 생성하였을 때나 기존에 작성한 DTD 파일을 읽어 들였을 때 사용자가 삽입, 삭제 또는 수정할 때 사용되는 모듈이다.

그림 5는 편집 모듈의 구성도이다. 그림 5의 구성도를 보면 DTD 트리에서 사용자가 마우스 이벤트를 발생시키게 되면 편집기능을 선택하는 메뉴를 출력하여 사용자가 원하는 편집기능을 선택한다. 편집기능을 선택하면 DTD 트리의 구조 정보를 갱신하여 DTD 트리를 사용자에게 다시 보여지게 된다.

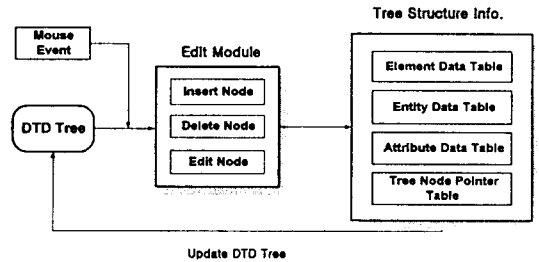


그림 5 DTD Tree 편집 모듈의 구성도

(1) 삽입 기능

삽입은 트리의 한 노드에서 노드를 형제(sibling) 또는 자식(child) 관계로 삽입하여 엘리먼트의 구조를 새롭게 정의할 때 사용되는 기능이다.

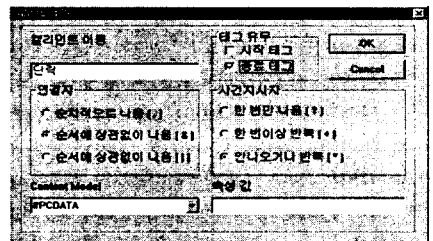


그림 6 엘리먼트 추가 다이얼로그 박스

DTD 트리에서 사용자가 엘리먼트를 삽입할 때 마우스 이벤트에 의해 삽입 메뉴를 선택하게 되는데 사용자는 삽입하고 싶은 노드에서 마우스 이벤트를 발생한다. 마우스 이벤트가 발생하면 삽입을 위한 다이얼로그 박스가 출력되어 사용자가 원하는 값을 입력하여 삽입을 하게 된다. 그림 6은 엘리먼트 삽입을 위한 다이얼로그 박스이다.

(2) 삭제 기능

삭제는 트리의 한 노드를 삭제하여 더 이상 필요 없게 된 엘리먼트를 지우기 위한 기능으로 사용자가 트리에서 마우스 이벤트에 의해서 선택된 트리 노드를 삭제한다. 삭제를 선택한 노드는 더 이상 유효하지 않으므로 화면의 트리 상에서 제거하게 된다. 삭제를 선택한 노드에 자식 노드가 있으면 자식 노드는 모두 유효하지 않으므로 자식 노드를 모두 삭제되게 설계하였다.

(3) 수정 기능

수정 기능은 이미 만들어진 엘리먼트에서 이름이나, 태그 유무, 연결자, 사건 지시자 등을 사용자가 재정의할 때, 수정 하고자 하는 노드로 이동하여 마우스 이벤트를 발생 시켜, 수정 대화상자에서 재 정의를 하여 엘리먼트의 내부 데이터 구조를 변경할 때 사용되는 기능이다.

3.5 DTD 생성 모듈

DTD 생성 모듈은 사용자가 편집기능 즉 삽입, 삭제, 수정 등의 기능으로 DTD 트리를 만들고 DTD 트리 생성 모듈에서 만들어진 DTD 트리를 다른 SGML 편집기나 브라우저에서 사용 가능하도록 *.dtd 파일로 만들 수 있게 메모리에 각 엔티티, 엘리먼트, 어트리뷰트 리스트를 저장하여 DTD 해석모듈에게 넘겨주는 기능을 한다. DTD생성 모듈에서 만들어진 엔티티, 엘리먼트, 어트리뷰트 리스트들을 DTD 해석 모듈에서 SGML DTD 규칙에 맞는지 오류 검사를 한 후 최종적으로 ASCII 파일로 저장한다.

3.6 DTD 생성기 실행의 예

본 논문에서 개발한 DTD 생성기를 위해 사용된 시스템 구현환경은 다음과 같다.

운영체제로는 WINDOWS 98 환경 하에서 개발 도구로는 VISUAL C++ 5.0, MFC 4.2를 사용하였다.

그림 7은 본 논문에서 구현한 SGML DTD 생성기의 실행화면을 가시화 한 예이다.

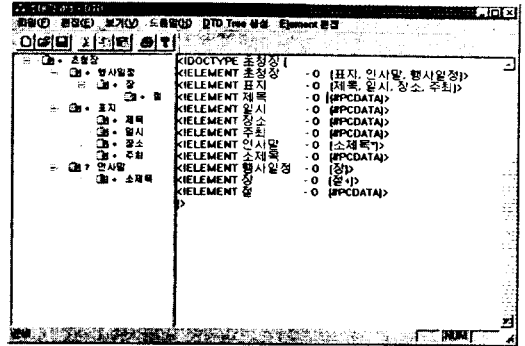


그림 7 DTD 생성기의 실행화면을 가시화한 예

IV. 결론

오늘날 WWW 문서 표준인 HTML은 모든 멀티미디어 문서를 표현하기에 많은 제약이 있으며, 이에 XML 이 새로운 WWW 문서의 표준으로 자리 잡을 것이 확실시되고 있다. 이는 WWW의 문서 표준이 SGML 형태로 전환되어 감을 의미하고 있다.

이와 같이 SGML의 중요성이 부각됨에 따라 본 논문에서는 SGML 문서의 구조를 정의하는 DTD 생성기를 설계 및 구현하였다. 현재 개발된 DTD 생성기는 대부분 외국에서 제작되어 한글을 완벽하게 지원하지 못하는 문제가 있고, 대부분 SGML DTD 작성 규칙을 DTD 개발자가 모두 알고 있어야 개발할 수 있는 단점을 가지고 있다.

이에 따라 본 논문에서는 SGML에서 규정하는 DTD 작성 규칙에 따라 DTD를 생성 및 편집할 수 있고, 한글 처리가 가능한 DTD 생성기를 설계 및 구현하였다. DTD는 문서의 논리적 구조를 파악하기 어렵기 때문에 문서의 논리적 구조를 파악하기 쉽게 트리 형식으로 설계 하였고, 사용자가 DTD의 작성 규칙을 자세히 모르더라도 GUI(Graphic User Interface)를 기반으로 하여 쉽게 DTD 트리를 작성하는 것으로 DTD를 작성할 수 있도록 구현하였다. DTD를 작성한 뒤 다른 SGML 응용 프로그램에서 사용이 가능하도록 SGML 형식으로 DTD 파일을 생성하도록 하였으며, 이미 생성된 DTD 파일을 불러서 편집이 가능하다.

본 논문에서 설계 및 구현한 DTD 생성기는 DTD의 논리 구조를 쉽게 파악할 수 있어서 SGML 문서 작성자가 구조 보다는 내용에 더 많은 시간을 할애하여 양질의 SGML 문서를 작성할 수 있고 다양한 DTD 구축으로 크게는 전자도서관 및, 학술 논문 DB 구축등 새로운 서비스 응용이 용이하고, 문서 양식의 배포는 DTD의 배포로 단순화 되어 문서 양식을 공유 할 수

있고, 네트워크 환경에서는 SGML 브라우저로 문서를 서로 공유할 수 있는 효과를 기대할 수 있다.

참고문헌

- [1] "Introducing the Standard Generalized Markup Language", Ominicon, open system
- [2] Eric van Herwijnen, "PRACTICAL SGML SECOND EDITION", KLUWER ACADEMIC PUBLISHERS, 1994
- [3] ISO 8879 (Standard Generalized Markup Language), 2차 수정본, 1986
- [4] Martin Colby & David S, Jackson, "Special Edition Using SGML", Que 1996
- [5] 고승규, "SGML 응용 시스템을 위한 SGML 파서 및 API의 설계", 연세대학교 석사학위 논문, 1993
- [6] 강치원, "SGML DTD 작성을 위한 대화형 편집 시스템에 관한 연구", 광운대학교 석사학위 논문, 1994
- [7] 김택근, "X 윈도우 상에서 SGML DTD 자동 생성 시스템 설계 및 구현", 배재대학교 석사학위 논문, 1996
- [8] 정희경, 현득창, 이수연, "SGML 가이드", 사이버 출판사 1998
- [9] 한상훈, 정신규, 김선우, 황대훈, "네트워크 상의 멀티미디어 정보 공유를 위한 SGML 브라우저의 설계 및 구현", 한국 정보 처리 학회 춘계 학술 발표 논문집, 1998
- [10] 정신규, 한상훈, 이종석, 황대훈, "문법 오류검사 기능을 갖는 SGML 문서 편집기의 설계 및 구현", 한국 정보 처리 학회 추계 학술 발표 논문집, 1998