

# ISO DSSSL을 기반으로 한 SGML 웹 브라우저

오박한주\*, 서상수\*\*, 이관진\*\*\*, 오병균\*  
\*목포대학교 정보공학부, \*\*(주)넥스텍  
\*\*\*초당대학교 컴퓨터학과

## SGML Web Browser based on DSSSL of ISO

Han-Ju Park\*, Sang-Su Seo\*\*, Pal-Jin Lee\*\*\*, Byeong-Kyun Oh\*  
\*School of Information Engineering, Mokpo National University  
\*\*Nextech.Co.Ltd., \*\*\*Dept. of Computer Science, Chodang University

### 요 약

ISO W3C에서 제안하고 있는 SGML은 문서의 논리적인 구조만을 기술하는 표준이기 때문에 레이아웃이나 프리젠테이션(Presentation)에 관한 정보는 포함하지 않는다. 따라서 SGML 문서를 보다 더 효율적으로 프리젠테이션(또는 브라우징)하기 위해서는 이에 대한 정보인 스타일시트(stylesheer)를 별도로 제공하여야 한다. 본 연구는 서로 다른 저작환경으로부터 구성된 문서를 사용자가 투명하게 뷰잉 또는 브라우징 할 수 ISO DSSSL 기반의 SGML 웹 브라우저를 이용하여 설계 및 구현한다.

### 1. 서론

인터넷상의 주요 문서가 HTML(Hyper Text Markup Language) 기반의 문서 활용에서 SGML(Standard Generalized Markup Language) 또는 XML(eXtensible Markup Language) 문서 포맷 영역으로 확장되어 가고 있기 때문에 기존의 정보관리 및 접근방식에도 새로운 인식이 필요한 상황이다. 따라서 멀티미디어 정보관리시스템은 이러한 국제적인 추세를 충분히 고려함으로써 네트워크 상에 분산된 데이터를 일관성있게 관리할 수 있어야 하며, 사용자에게는 이러한 정보들이 투명하게 제공되어야 한다.

현재 국·내외적으로 SGML 문서를 브라우징할 수 있는 다수의 프로덕트(Product)가 개발되어 있다[3,6,11,13]. 이들 SGML 브라우저는 구조화된 정보의 재사용에 상당한 융통성을 제공하기는 했지만 모든 문서의 레이아웃 및 포맷팅(Layout and Formatting)에 독자적인 기술을 사용하고 있다. 이것은 국제표준화기구인 ISO W3C(WWW Consortium)에서 주장하고 있는 개방형 문서표준인 SGML의 본래 취지에 어긋나며, 또한 시스템간의 문서 상호호환을 어렵게 만드는 요인이다.

ISO W3C에서 제안하고 있는 SGML은 문서의 논리적인 구조만을 기술하는 표준이기 때문에 레이아웃이나 프리젠테이션(Presentation)에 관한 정보는 포함하지 않는다[7]. 그러므로 SGML 문서를 프리젠테이션하기 위해서는 이에 대한 정보인 스타일시트(stylesheer)를 별도로 제공하여야 한다. 이를 위해서 ISO에서는 SGML문서의 다양한 처리정보를 기술하기 위해 표준안을 발표하였는데, 그 대표적인 것이 FOSI(Format Output Specification Instance)와 DSSSL(Document Style Semantics and Specification Language)이다. FOSI는 출력 사양에 대한 표준을 얻으려는 초기 시도로서 그 기능적인 측면에서 많은 제한점이 있다. 반면에 DSSSL은 ISO 10179에서 발표한 표준으로 SGML 문서 인스턴스(Document Instance :

DI)내의 특정 내용과 그 내용에 기초한 출력 포맷팅 구조를 인식할 수 있는 기능을 포함하기 때문에 FOSI에 비해 더 많은 융통성과 기능을 제공하고 있다[9,10].

본 논문은 서로 다른 저작환경으로부터 구성된 문서를 사용자가 투명하게 뷰잉 또는 브라우징할 수 ISO DSSSL 기반의 SGML 웹 브라우저를 설계 및 구현한다. 여기에 사용된 개발 환경은 Grif 군(Family) 도구를 이용하였다. 특히 이 틀은 구조적인 문서 표준화와 호환성을 지니고 있고, 또한 국내의적으로 많이 사용되고 있는 상황이다.

본 논문의 순서는 1장에 이어 2장에서는 웹 상의 문서를 HTTP에 의해 가져올 때 필요한 요구사항들을 정리하고, 3장에서는 웹 브라우저에 대한 구조를 다루었다. 4장에서는 설계된 웹 브라우저를 Grif 환경을 이용하여 구현하고, 마지막으로 5장에서는 결론을 맺었다.

### 2. 요구사항

오늘날 웹 브라우저는 단순한 스케줄을 따라 일련의 작업을 수행한다. 웹 브라우저는 태그용 HTML 문서를 스캔하고, 일반 ID(태그)를 읽고, 태그를 프로세스하고, 다음 태그를 찾는다.

SGML 기능을 웹을 통해 확장하는 목적은 브라우저가 DTD를 load해서 각 문서에 독립적으로 규칙을 프로세싱하도록 하는데 있다. DTD와 스타일시트는 브라우저가 서버에서 문서를 가져올 때 동적으로 load 되어야 하며, 다음과 같은 요구사항들이 필요하다.

- ① 모든 스타일시트를 동일한 종류의 mechanism으로 읽을 수 있게 스타일시트를 정의하는 공통 방법을 만든다.
- ② HTTP 서버로 하여금 각 문서의 DTD나 스타일시트 정보를 사용 가능하게 유지하도록 한다.
- ③ 각 HTTP 서버로 하여금 완벽한 태깅으로 SGML 문서를 확인하도록 한다.
- ④ 별도의 정보를 쉽게 처리하도록 HTTP 헤더 마크업을

표준화한다.

⑤ HTTP 서버가 이미 실행한 내용을 또는 파싱을 실행해 선 안되는 내용을 중복하지 못하도록 모든 브라우저를 영리한 파서로 만든다.

⑥ 모든 브라우저가 완전한 SGML 양방향 링크(외부 브라우저 시나리오일 경우, SGML 브라우저가 HTML 브라우저에게 다시 말해야 하는)를 지원하게 만든다.

### 3. SGML 웹 브라우저 설계

저작 환경에 의해 생성된 문서를 사용자가 투명하게 볼 수 있는 웹 브라우저의 기본 구조는 그림 1과 같다.

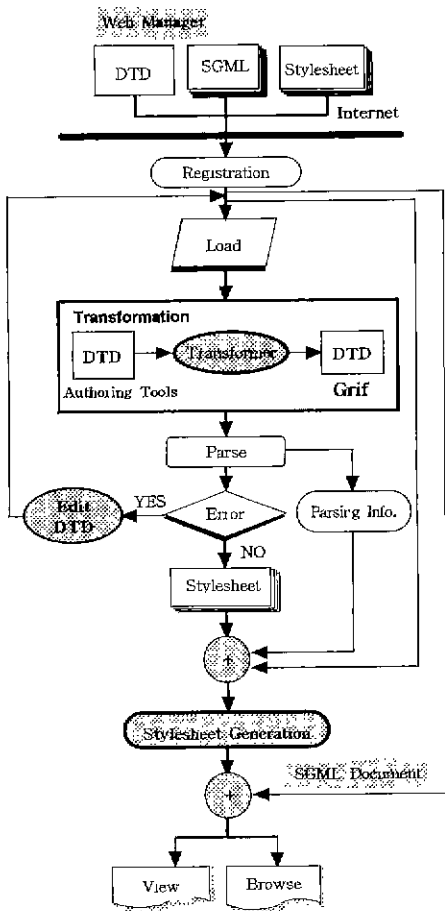


그림 1. 웹 브라우저 구조

SGML 브라우저 요구사항에 따라 웹 상으로부터 SGML 문서를 가져올 때 DTD, SGML 문서, 그리고 스타일 정보가 local memory에 등록된다. 다음에 Grif에 있는 변환 기능을 이용하여 웹 상에서 가져온 DTD를 Grif DTD로 변환하고, Builder를 이용하여 파싱한다. 에러가 없으면 파싱 정보를 이용한 스타일시트를 생성하고 웹 상에서 가져온 SGML 문서를 브라우저한다. 만일 에러가 발생하면 DTD를 편집하여 다

시 파싱을 수행한다. 그러나 DTD만을 변환하여 스타일시트 정보를 생성하고, 이를 이용하여 SGML 문서를 브라우저할 경우 기존의 프리젠테이션 정보와의 많은 차이가 있기 때문에 여기에 파싱 정보와 웹 상에서 가져온 스타일시트 정보를 이용하여 SGML 문서를 브라우저한다.

### 4. SGML 웹 브라우저 구현

본 논문에서는 설계된 SGML 웹 브라우저의 구현은 Grif 도구와 Windows OS 하에서 Visual C++6.0을 이용하여 구현하였다. Grif는 I4I에서 개발된 SGML 소프트웨어로 현재 국·내외에서 가장 많이 쓰이고 있는 툴이다.[5] 특히 Grif가 많이 사용되고 있는 이유는 P(Presentation)라고 하는 자체 스타일시트를 포함하고 있는데, 이것의 특징은 ISO의 DSSSL에서 제안하고 있는 거의 모든 기능을 포함하고 있다는 것이다. 따라서 DSSSL을 이용한 스타일시트 변환을 하지 않더라도 곧 바로 스타일시트를 생성할 수 있다는 것이다. 또 하나는 GATE 라는 엔진을 가지고 있어 어떠한 기능도 Grif 제공하고 있는 API를 이용하여 개발할 수 있다는 것이다.

#### 4.1 Grif SGML Editor 통합 환경 갱신

Grif SGML Editor 통합 환경에 추가될 기능은 문서 변환 메뉴와, 현재 별개로 분리되어 있는 Application Builder, 그리고 DTD 편집 및 스타일시트 정보를 편집할 수 있는 텍스트 에디터 3 가지이다. 이와 같은 새로운 기능의 추가는 Grif 제품군 중의 하나인 GATE를 이용하여 수행된다.

#### 가. GATE

GATE는 Grif Application Toolkit Environment의 약어로서 문서 데이터베이스 시스템이나 워크플로우 소프트웨어와 같은 외부 응용 프로그램들을 쉽게 접목시킬 수 있다. 또한 SGML 편집기에 새로운 메뉴 또는 새로운 함수를 쉽게 추가할 수 있기 때문에 DTD에 대한 새로운 문서가 주어질 경우 사용자의 요구사항에 맞도록 재구성할 수 있는 환경을 제공해 준다.

#### 나. Interface 언어

Interface (I)언어는 Message와 Action을 이용하여 편집기 상에서 사용자의 인터페이스를 최적화시킬 수 있는 도구이다. Message는 주어진 분야에 적용될 행위의 유형(또는 편집기능)을 나타내며, Action은 사용자 응용 코드로서 텍스트 기반의 메시지를 구현한다.

GATE 환경을 이용한 사용자 인터페이스 개발은 다음과 같은 순서로 수행된다. 먼저 추가하고자 하는 메뉴 항목을 정의하고, 이 메뉴 항목을 인터페이스 언어 구성 형식에 따라 message와 action을 기술하여 임의의 파일(예 sample.i)로 저장한다. 다음으로 Grif 군의 개발환경으로 지원하고 있는 그림2의 Builder를 이용하여 인터페이스 파일을 컴파일하면 3가지의 파일(sample.h, itf.c, sample.csk)이 생성된다

#### 4.2 문서 브라우징

문서 브라우징을 위해 초기에 나타나는 Grif SGML Editor의 통합 환경은 문서와 도움말로 구성된다. 문서(F)는 세글

(N), 불러오기(O), 그리고 끝으로 이루어져 있으며, 도움말(H)은 Grif Editor 운영에 따른 각종 도움말을 제공한다.

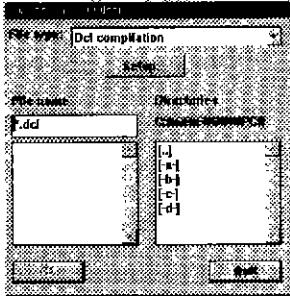


그림 2. Application Builder

가. SGML 문서작성

그림 3은 주어진 DTD가 chapter인 경우 이에 대한 새로운 SGML 문서 작성 준비 상태를 보여준다. 문서는 에디터 막스 상단에 있는 엘리먼트를 이용하여 작성 가능하며, 필요에 따라 각 엘리먼트별 스타일시트 정보 생성이 가능하다. 이 기능은 문서 변환 과정을 거친 후 Application Builder를 이용하면 DTD에 대한 기본적인 스타일시트 생성이 가능할데, 여기에 엘리먼트별 스타일시트 정보를 생성할 경우 주어진 문서에 대한 최종적인 스타일정보 생성 과정은 끝나게 된다.

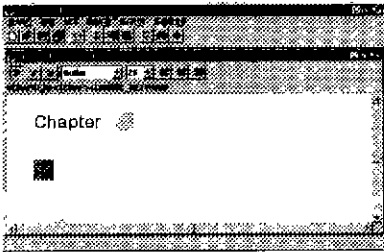


그림 3. 새로운 SGML 문서 작성

나. SGML 문서 불러오기

SGML 문서 불러오기는 인터넷과 local memory 두 가지로 분류하여 SGML 문서를 불러올 수 있는데, 설계 및 구현된 SGML 웹 브라우저를 이용하여 쉽게 처리할 수 있다. Grif SGML Editor를 이용하여 문서를 불러올 경우 가장 먼저 그림 4와 같은 HTTP 지원의 웹 주소 대화상자가 화면상에 나타난다.



그림 4. SGML 문서 불러오기

Address란에 주어지는 부분은 HTTP(HyperText Transfer Protocol) 또는 FTP(File Transfer Protocol)에 의해 지원되는 인터넷상의 SGML 문서에 대한 주소 또는 local memory에 있는 파일에 있는 주소가 된다.

그림 5는 브라우저 될 파일을 지정하여 읽어들이는 경우 화면에 뷰잉된 결과를 보여주는데, 만일 SGML 문서가 DTD에 대한 chapter일 경우 이에 대한 문서의 출력을 보여주고 있다.

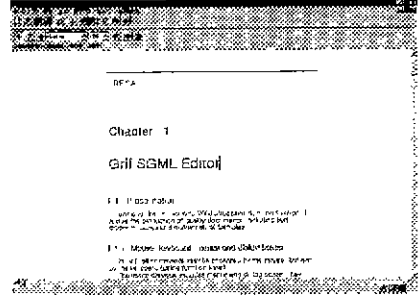


그림 5. 문서 브라우징

5. 결론

네트워크 상에서 서로 다른 SGML 문서를 서로 공유하기 위해서는 분산된 데이터를 일관성있게 관리하여야 하며, 사용자에게는 투명하게 제공되어야 한다. 이를 위한 다양한 처리 기술(CSS, FOSI, DSSSL, XSL)들이 등장하고 있다.

본 논문에서는 이를 해결하기 위해 Grif 개발 환경을 이용하여 다양한 문서를 뷰잉할 수 있는 DSSSL 기반의 SGML 웹 브라우저를 설계 및 구현하였다. 개발된 웹 브라우저는 서로 다른 저작환경으로부터 구성된 SGML 문서를 사용자에게 일관성 있고 투명하게 프리젠테이션 할 수 있을 뿐만 아니라 SGML로 기술된 전자문서로부터 사용자가 필요한 문서를 보다 쉽게 획득할 수 있는 시스템이다.

향 후 본 논문의 결과를 토대로 범용 XML 문서 웹 브라우징에 대한 연구를 지속적으로 수행할 것이다.

참고문헌

- 1) <http://swallow.yonsei.ac.kr/~sgml/pijabs/>
- 2) \_\_\_\_, "Web 브라우저 개발 지침서," 한국전산원, 1996.11.
- 3) Sean McGrath, "PARSEME.1ST : SGML for Software Developers, Charles F.Goldfarb Series On Open Information Management, 1998.
- 4) Martin Colby & David S. Jackson, 김숙자 옮김, Special Edition SGML, 성안당, 1997.
- 5) Grif Products Manual, idi, France, 1998.
- 6) <http://www.SoftQuad.com/resources/sgml/index.html>
- 7) International Standard of ISO/IEC, Information Technology-Process Language-Docment Style Semantics and Specification Language(DSSSL), ISO/IEC 10179:1996(E)
- 8) \_\_\_\_, "SP," <http://www.jclark.com/sp/>
- 9) Paul Prescod, "Introduction to DSSSL," <http://itrc.uwaterloo.ca:80/~papresco/dsssl/tutorial.html>
- 10) Jade-James, "DSSSL Engine," <http://www.jclark.com/jade/>
- 11) <http://collie.fujitsu.com/hybrick/>
- 12) <http://www.oasis-open.org/cover/free-pan.html>
- 13) <http://www.oclc.org:5046/oclc/research/panorama/panorama.html>