

XML 기반의 Shopping Agent System 설계 및 구현

서인규⁰ 최병석 한성국

원광대학교 컴퓨터공학과

(inca, skhan}@wonkwang.ac.kr, bschoi@kunjang.ac.kr

Design and Implementation of Shopping Agent System based on XML Document

In-Kyou Seo⁰ Byung-Soek Choi Sung-Kook Han

Dept. of Computer Science, Wonkwang University

요 약

웹의 표준은 XML을 기반으로 개발, 발전 되고 있는 추세이며, 이러한 XML 기술은 전자상거래, 문서 변환, ERP등의 응용 분야에 활발하게 적용되고 있다. 이와 더불어 XML 문서화에 따른 데이터 처리 기법도 중요하게 되었다. 데이터 처리 기법에 에이전트 기술을 이용함으로써 불필요한 HTML 문서 내용을 처리하는 기존의 방법보다 구조화된 데이터를 검색하여 결과의 정확성과 신속성을 가질 수 있다. 본 논문에서는 이러한 논리적인 XML 문서를 기반으로 한 쇼핑몰을 구축하고 그에 대한 데이터를 수집, 처리하는 에이전트를 설계, 구현하였다.

1. 서론

최근 컴퓨터 정보 통신 기술의 발달로 인한 정보의 양이 급증함에 따라 문서 처리의 중요성이 증가하고 있다. 이러한 정보 환경의 변화는 쉽고 간편함을 제공하지만 데이터들은 과다하게 증가함으로써 복잡한 형태를 지닌다. 이에 따른 복잡한 문서를 처리하기 위한 노력으로 전자출판, 디지털 도서관, 전자 상거래등과 같은 대량의 문서를 구조화 하여 관리하거나 표준화된 방법으로 처리하기 위한 대안으로 XML이 요구되고 있다 [9].

XML 문서는 구조화 되어 있기 때문에 이를 사용할 경우 많은 응용 프로그램에 대한 표현력의 간결성 및 데이터 검색이나 표현에 있어서 많은 이점을 가진다 [6, 7, 8, 10].

많은 양의 문서를 실시간으로 처리해야 하는 전자 상거래나 쇼핑몰 시스템 같은 경우 정보를 처리하는 능력을 자율적으로 대신하여 주는 에이전트 기법이 많이 사용되고 있는 추세이다. 이러한 시스템들의 문서나 정보들이 XML문서로 구조화 되어감에 따라 에이전트를 이용한 XML 문서를 처리하는 방법이 요구된다 [5, 8, 9, 10].

현재 대다수 에이전트의 경우 3-계층(3-Tier)구조의 데이터 베이스 시스템이나 웹 문서로부터 정보를 추출하고 수집한다[3, 4]. 수집된 결과들은 HTML 태그(Tag)들을 통해 변형된 후 클라이언트의 웹 브라우저에 보여진다. 그러나 문서가 데이터 타입과 같은 원본 그대로의 정보를 가지고 있는 XML 문서로 구조화 되어 있다면 기존의 시스템보다 데이터 처리 방법이 간단한 에이전트를 설계할 수 있다. 각 응용 시스템은 DTD를 참조하여 이루어져 있기 때문에 본 논문에서는 이 DTD 정보를 기반으

로 한 XML 기반 쇼핑몰 에이전트 시스템을 설계 구현하였다.

2. XML문서에 기반 한 쇼핑 에이전트 시스템의 특징

분산 웹 어플리케이션을 구현할 때 가장 많이 쓰는 3-계층 모델은 각각의 모듈이 다른 곳에서 개발되어도 응용하는 데는 어려움이 없다[8]. 3-계층 중 데이터 베이스 계층은 데이터를 저장하는 곳으로 구조화 된 정보, 즉 데이터 베이스나 XML문서와 같은 어떠한 구조로도 구성할 수 있으며 논리적으로 구조화 되어 있기 때문에 다른 계층 모듈에 비해 활용도가 넓다. 본 논문에서 제안한 쇼핑 에이전트는 다음과 같은 특징을 지니고 있다.

- 1) 간결성(Simplicity): 대다수의 쇼핑몰들이 검증된 DTD를 사용하게 된다면 어떤 종류의 쇼핑몰이든 이러한 DTD를 토대로 하여 각각의 목적에 맞는 웹 프로그램을 개발할 수 있다. 그렇게 된다면 DTD 구조와 주소만 알아도 쇼핑몰의 데이터에 접근을 가능하게 하여 검색 에이전트를 쉽게 구축하여 이전의 사용하던 방법보다 데이터 접근을 간결하고 원할 하게 할 수 있다.
- 2) 자료구조의 용이: 쇼핑몰이 검증된 DTD를 참고하여 개발이 된다면 그 데이터들은 XML의 요소인 논리적 구조를 가진 자료 구조로써 데이터가 여러 곳에 분산되어 있다고 해도 엘리먼트(Element)의 조합, 삭제, 갱신, 결합 등을 간단하게 할 수 있다. XML은 복잡한 자료구조를 표현 하는데 있어서 충분한 능력을 가지고 있으므로 본 논문에서 구현한 에이전트 시스템에

서도 이를 활용, 응용하고 있다.

3) 문자 처리: XML은 유니코드를 지원하기 때문에 거의 모든 문자를 처리, 표현 할 수 있다. 쇼핑물이 어떠한 시스템으로 이루어져 있다고 해도 처리할 수 있다. DTD를 기반으로 한 XML 쇼핑물들을 에이전트가 검색하는 경우 트리 구조의 자료 접근과 기존의 이기종 시스템에 HTML을 포함하는 데이터 접근 방법이나 각각 상이한 DBMS의 접근 방법보다 지능적이고 간단할 수 있다.

3.XML기반 쇼핑 에이전트 설계 및 구조

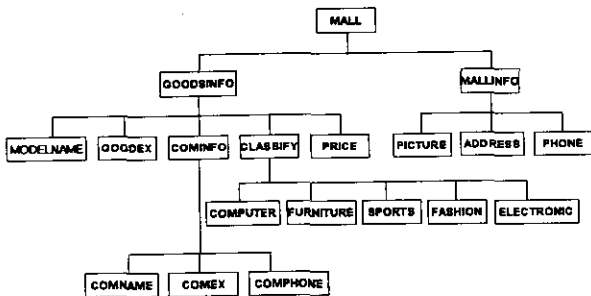
에이전트는 다양한 플랫폼에서 정보를 탐색하고 추출해야 하기 때문에 이기종 간의 환경을 동일하게 지원하는데 언어가 필요하다. 본 논문에서는 에이전트를 개발함에 있어서 플랫폼에 독립적, 안정성, 생산성 그리고 이동성 즉, XML 문서를 처리하는데 가장 접근이 용이하고 많은 API가 제공되는 Java를 이용했다. 또한 XML을 처리하기 위해 IBM에서 제공되는 XML for JAVA API를 이용하여 설계 했다.

이기종간의 분산 환경에서 효율적으로 다른 언어로 쓰여진 에이전트의 운용 환경을 제공하기 위해서 OMG의 CORBA 인터페이스인 MASIF 인터페이스를 사용했다.

XML 데이터를 처리하는데 있어서 IBM의 XML for JAVA는 처리기로서 파싱과 생성의 두 가지 주요 기능을 가지는 API를 제공한다.

3.1 시스템의 구조

XML기반의 쇼핑물을 검색 할 경우 DTD를 기반으로 한 트리 형태의 구조로 정보가 이루어져 있기 때문에 이 DTD를 접근, 제어 할 수 있는 DOM API를 이용하여 문서를 갱신, 추가, 삭제, 재결합 할 수 있는 기반을 제공한다. (그림 1)은 쇼핑물 DTD와 DOM트리 구조를 나타낸 그림이다.

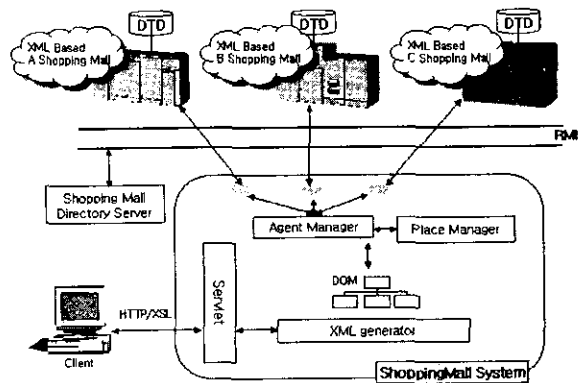


(그림 1) DTD 트리 구조

```
<?ELEMENT MALL (GOODSINFO, MALLINFO)>
<ELEMENT GOODSINFO (MODELNAME=, GOODEX=, COMINFO=, CLASSIFY=, PRICE=)
<ELEMENT MODELNAME (#PCDATA)>
<ELEMENT GOODEX (#PCDATA)>
<ELEMENT PRICE (#PCDATA)>
<ELEMENT COMINFO (COMNAME=, COMEX=, COMPHONE=)
<ELEMENT COMNAME (#PCDATA)>
<ELEMENT COMEX (#PCDATA)>
<ELEMENT COMPHONE (#PCDATA)>
<ELEMENT CLASSIFY (#PCDATA | COMPUTER | FURNITURE | SPORTS | FASHION | ELECTRONIC)
<ELEMENT PRICE (#PCDATA)>
<ELEMENT MALLINFO (PICTURE=, ADDRESS=, PHONE=)
<ELEMENT PICTURE (#PCDATA)>
<#ATTLIST Picture
File CDATA #REQUIRED
>
<ELEMENT ADDRESS (#PCDATA)>
<ELEMENT PHONE (#PCDATA)>
<ELEMENT COMPHONE (#PCDATA)>
```

(그림 2) 쇼핑물 구성을 나타내는 DTD

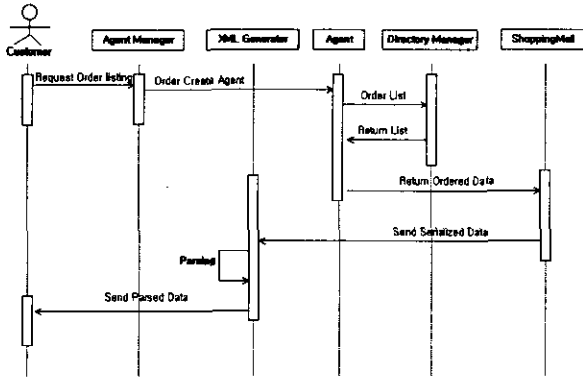
여러개의 쇼핑물들을 비교 검색 할 경우 데이터들은 DBMS 자체내의 데이터에서 HTML로 보여지는 정보들, 각 벤더들이 각각의 시스템에 별도로 정보를 관리하고 있는 것이다. 검색 에이전트는 각 쇼핑물에 대한 모든 데이터 구조를 알고 있어야 하고 여러 데이터들을 조합하여 문서를 추출할 경우 상당한 노력이 필요하다. 또한 인터넷 상의 메타 데이터에 대하여 문서를 추출할 경우 많은 인력과 시간을 요구하며 검색을 통한 정보 활용 측면에 있어서도 데이터의 정확도가 많이 떨어진다. 그러나 XML문서로 모든 데이터를 구조화하면 DTD의 정보만을 참조함으로써 쇼핑물 정보는 XML/DOM 형태로 접근이 가능하므로 에이전트는 클라이언트가 원하는 각각의 쇼핑물에 대해서 문서의 요소(Element) 내의 정보를 단계적으로 검색해서 클라이언트에게 특정 정보만 빠르고 정확하게 전달 할 수 있는 것이다. 전달된 결과물은 클라이언트의 종류에 따라서 HTML이나 다른 형태로 반환된다. (그림 3)은 본 논문에서 구현한 시스템의 구조도이다.



(그림 3) 시스템 구조도

XML과 HTML은 발전하고 있기 때문에 XML을 처리하는 부분과 HTML을 생성하는 부분을 분리하여 처리 하였으며 객체 지렬화를 하기 위한 데이터층을 중간에 삽입 하였다. 동작 과정을 살펴 보면 클라이언트가 어떠한 물건과 가격 정보를 입력 하였을 경우 XML 문서를 생성하기 위해 파서는 구문분석을 마친 후 DOM 트리를 생성하여

생성기로 전달을 한다. 에이전트는 이 이벤트를 객체를 받아 데이터를 처리한다. 하나의 상품 데이터는 회사이름, 모델 번호, 모델 이름, 가격등의 정보 데이터로 이루어진다. 에이전트에 의해 처리된 데이터는 객체 직렬화한 후 XML 엘리먼트와 속성을 참고하여 DOM 트리에 추가한다. DataConverter 객체는 Shopping-Mall.dtd 파일에 따라 노드의 값들을 HTML로 변환한다.(그림 4)는 본 논문에서 제안한 시스템의 이벤트 흐름 과정이다.



(그림 4) 이벤트 흐름

에이전트는 현재 개발 되어 있는 비교 쇼핑몰 에이전트의 동작 방법과 비슷한 에이전트 생성과정을 수행한다. CORBA 메커니즘인 디렉토리 매니저 인터페이스를 거친 후 해당 사이트의 협상 에이전트와 협상을 한 후 여러 사이트에서 조합한 결과물들을 출력을 담당하는 OutGenerator 객체로 전송하여 클라이언트에게 HTML 형식으로 결과를 출력하는 구조를 가지고 있다.

(그림 4)는 본 논문에서 구현한 3 개의 임의의 XML 문서로 이루어진 사이트를 에이전트가 DTD를 참고한 엘리먼트 정보를 요구함으로써 데이터를 구조적으로 검색하는 모습을 보인 그림이다.

에이전트는 디렉토리 매니저에 등록되어 있는 인터페이스를 거친 후 3 개의 사이트를 검색하여 결과물 을 찾아 내었다. 쇼핑몰 에이전트 시스템은 데이터가 파서를 거쳐 변환되기 때문에 속도면에서는 다른 시스템의 처리 과정에 비해 높지 않은 상태이다. 하지만 자연어 처리 방법, 메타 데이터 검색 그리고 에이전트 분야 등에 있어서 구조적 문서의 데이터 검색으로 활용할 수 있는 여러 가지 측면은 기존의 방법 보다 간단하고 구조화에 따른 데이터 추출의 정확도가 높다는 이점이 있다.

4. 결론

컴퓨터를 기반으로 한 전자 문서 처리와 이기종 시스템간의 데이터 교환이 더욱 방대해짐에 따라 이를 처리하고 표현하는 방법은 다양해져 가고 있다. XML은 메타 언어뿐만 아니라 데이터로서 사용이 가능하며 각 내용이 구조화 됨에 따라 새로운 시스템 응용을 원활하게 할 수 있다. 이에 웹의 표준은 XML을 기반으로 개발, 발전 되고 있는 추세이며 또한 XML 기술 개발이 표준 개발 기구의 독자적인 주도보다는 주요 소프트웨어 벤더들이 XML

을 적극적으로 지원하고 그 발전도 동시에 이루어 지고 있음에 따라 기반 프레임워크 위에 XML을 어떤 형태로 사용할 것인가에 대한 다양한 어플리케이션 개발 방법론이 대두 되고 있다. XML기술은 특히 전자상거래, 무선 인터넷, 문서 변환, ERP등의 분야에 응용분야가 활발하게 사용 될 것이며 이러한 분야에서 이루어지는 메타 데이터에 대한 검색 방법중의 하나인 XML문서를 다루는 에이전트 기법이 필요할 것이다. 본 논문에서는 비교 쇼핑몰 시스템에 대한 에이전트 기술이 적용 가능한 방법과 에이전트가 인터넷 기술의 표준으로 자리잡게 될 XML 문서를 다루는 방법에 대하여 연구하고 및 구현 하였다.

이 같은 XML 기술의 활용과 연구로 데이터 정보 추출을 위한 자연어 처리 기법에 대한 연구가 진행되어야 할 것이며 XML 응용 소프트웨어 분야가 아닌 XML 핵심기술에 대한 연구가 필요할 것이다. 또한 데이터 베이스 시스템에서의 XML 데이터의 효과적인 처리 방법 역시 중요하며 저장된 데이터를 어떤 형태의 DTD든지 그것을 따르는 XML 데이터 형태로 바꿔 만들 수 있는 기능등과 같은 연구가 필요할 것이다.

5. 참고문헌

[1] Dan Chang, Dan Harkey, " Client/Server Data Access with Java and XML," pp 106-130, 11, 1997
 [2] D. D. Roure et al. , " Agents for Distributed Multimedia information Management," Proc. of 1 In-t' 1 Conf. on Practical Applications of Intelligent Agents and Multi-Agents Technology, 4 , 1996
 [3] Elliotte Rusty Harold. " JAVA Network Programming?" , pp 302-340, 6, 1999
 [4] Frank Boumphrey, " Professional XML Applications." pp 208-252, 09, 1999
 [5] Hiroshi Maruyama, Kent Tamura, Naohiko Uromoto , "XML and Java Developing Web Applications", pp 282-311, 8, 2000.
 [6] Ivor Horton, " Beginning Java2" , pp 756-772 , 9 , 2000
 [7] Minar et al, "Cooperating Mobile Agents for Dynamic Network Routing." , MIT Media Lab,1999
 [8] 이강찬외 2 인, " XML 표준화 동향", 정보과학회지, 제 19 권 1 호, 제 140 호, 01, 2001
 [9] 이재호, " 에이전트 시스템의 연구 및 개발 동향" , 정보과학회지, 제 18권 제5호, 제 132호 ,5, 2000