

XML 기반의 정보 통합을 위한 OODB2XML 래퍼의 설계 및 구현*

김태현^o, 김경일, 이강찬, 이규철
충남대학교 컴퓨터공학과 데이터베이스 연구실

Design and Implementation of OODB2XML wrapper for XML-based Information Integration

Taehyun Kim^o, Kyungil Kim, Kangchan Lee, Kyuchul Lee
Department of Computer Engineering, Chungnam National University

요 약

W3C에서 제안한 XML 표준안이 발표된 이후 인터넷을 기반으로 하는 많은 다양한 분야에서 XML을 활용하고자 하는 연구들이 진행중이다. 그 중 한 분야가 인터넷상에 널리 퍼져있는 기존의 다양한 정보 시스템들을 XML을 매개로 하여 통합하고자 하는 연구이다. 즉 분산 환경에서의 각자 고유한 스키마를 가지고 운영되고 있는 기존의 정보시스템들을 XML이라는 공통 데이터 모델을 사용하여 표현함으로써 통합된 정보를 제공하는 것이다. 따라서, 이러한 연구에서는 기존의 정보시스템의 데이터에 대한 XML화된 질의 방법을 제공하는 것이 필수적이며 질의에 대한 결과 역시 XML로 표현되어야 한다.

본 논문에서는 이러한 관점에서 객체지향 데이터베이스인 바다-III 데이터베이스를 대상으로 XML에 대한 질의 언어인 XQL을 사용한 질의 시스템과 그 결과를 XML로 변환하는 변환 시스템의 설계 및 구현 방법에 대해 논하고자 한다.

1. 서론

XML(eXtensible Markup Language)[1]은 W3C(World Wide Web Consortium)에서 표준화한 웹을 기반으로 하는 구조화된 문서를 기술하는 방법에 대한 표준이다. 현재 XML은 차세대 인터넷 기술로 각광을 받고 있으며 XML 응용에 대한 많은 연구가 진행중이다. 그 중에서도 인터넷상에 산재되어 있는 이질적인 형태의 정보들을 통합하는데 XML을 이용하려는 노력이 있다[2].

분산되어있는 기존의 정보 시스템들은 각자 고유의 데이터 모델을 사용하여 데이터를 표현한다. 이러한 이질적인 데이터 모델로 표현된 정보들을 통합하기 위해서 하나의 중립적인 데이터 모델이 필요하며, 이 데이터 모델로서 XML을 이용하려는 것이다. 이에 따라, 기존 정보 시스템의 데이터모델을 XML 데이터 모델로 변환하는 래퍼(wrapper)가 필요로 하게 되었고, DB2XML이라는 이름으로 관계형 DBMS에 저장된 정보를 XML로 변환하여 보여주는 프로그램도 개발되어져 있다[3].

하지만 관계형 DBMS 대한 XML 변환 시스템은 관계형 DBMS에 저장된 데이터가 가지는 특성에 의해 단순히 테이블, 레코드, 필드간의 포함관계만을 표현함으로써 원래의 데이터간의 관계 및 의미를 표현하는데 한계가 있다.

본 논문은 바다-III DBMS를 대상으로 XQL[4] 형태의 질의를 사용하여 객체지향 DBMS에 저장되어 있는 정보들을 검색하고, 그 결과를 XML로 변환하여 출력하는 래퍼 시스템을 설계 및 구현하였다. 객체지향 DBMS에서의 데이터는 객체 단위로 처리되어지며, 하나의 영속 객체는 여러 개의 다른 영속 객체들을 포함할 수 있고, 이러한 포함 관계는 트리 형태의 구조를 갖는다. 이러한 트리 형태의 객체 관계는 XML 엘리먼트 트리 구조로 쉽게 표현할 수 있다는 장점을 가진다.

2. 관련연구

2.1 XQL

1998년 9월에 XML Working Group에 제안된 XQL(XML Query Language)은 XML문서의 엘리먼트와 내용을 가리키고, 여과시키는 표기법으로써 XSL pattern syntax를 확장하여 사용한다. XQL은 특정 엘리먼트를 가리키고, 노드를 검색하는데 있어서 간결하고 이해하기 쉬운 표기법을 제공한다.

XQL에서 사용하는 표기법은 URI 디렉토리 표기법을 본떠서 사용한다. 예를 들면, book 엘리먼트에 포함된 author 엘리먼트를 찾으려면, 다음과 같이

book/author

라고 표시하면 된다. 질의결과는 선택된 것들의 집합으로 나타난다.

* 본 논문은 소프트웨어연구센터의 핵심응용기술과제인 XML 저장/검색 및 분산 문서 시스템의 설계 및 구현(과제호:99-11-02-01-A-2)으로 수행되는 과제임.

2.2 바다-III OQL

바다-III OQL은 select-from-where 문장을 기반으로 한 질의어이다. 이 질의어는 ODMG-93 OQL을 기반으로 정보 검색 질의를 위하여 ISO8777표준에서의 FIND 사양을 융합한 구문 구조를 가지고 있어서, 하나의 질의에 데이터베이스 및 정보 검색 질의 조건을 자유롭게 섞어서 사용할 수 있다[5].

바다-III에서의 데이터베이스 질의 대상은 단일 클래스가 되며 그 클래스는 다른 클래스를 포함할 수 있다. 따라서 질의어는 다음과 같은 형식으로 제한되며

```

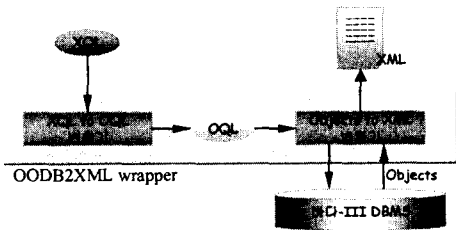
"select * from class-name where condition"
    
```

질의 결과는 from절의 class에 대한 영속 객체의 집합형태가 된다.

3. 시스템 설계 및 구현

3.1 System 개요

본 논문을 통해 구현된 래퍼 시스템은 사용자로부터 XQL 질의를 입력받고, 결과를 XML 문서로 넘겨줌으로써 사용자로 하여금 바다-III DBMS를 XML 저장소와 같이 사용할 수 있도록 한다. 하지만, 실제적인 데이터는 객체지향 DBMS에 저장되어 있는 객체들이며, 사용자가 질의한 XQL은 OQL로 변환되어 DBMS에 질의되고 결과로 나온 객체의 집합은 다시 XML문서로 변환되어 사용자에게 전달된다.



<그림 1> OODB2XML wrapper system 구조

3.2 XQL2OQL 변환기

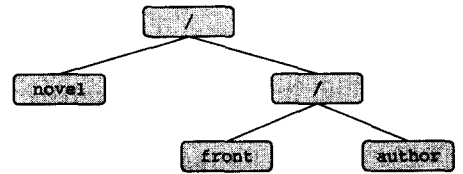
1) XQL2OQL 변환기의 개요

XQL2OQL 변환기는 사용자로부터 XQL을 입력받아서 객체지향 DBMS가 이해할 수 있는 OQL로 변환하는 모듈이다. XQL을 OQL로 변환할 때 XQL 표현식에서 나오는 각 엘리먼트들은 객체의 속성, 또는 그 객체가 가리키는 또 다른 객체로 매핑된다. 즉, XQL 표현 'novel/front/author'를 OQL로 변환하면, 'SELECT x.front.author FROM x in novel'로 나타난다.

2) XQL2OQL 변환기의 설계

입력받은 XQL을 OQL로 변환하기 위해서 XQL을 구조화된 형태로 바꾼 다음, 그 구조를 해석함으로써 OQL로 변환하게 된다. 그 구조는 파싱 트리라고 부르며, 이진트리의 형태를 가졌다. 예를 들어 XQL로 'novel/front/author'는 <그림 2>과 같은 파싱 트리로 바뀌게 되고, 이 파싱 트리를 해석하여 OQL로 변환하게

된다.



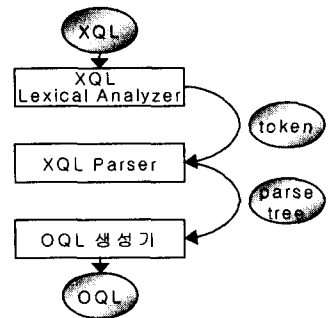
<그림 2> novel/front/author에 대한 파싱 트리

3) XQL2OQL 변환기의 구현

XQL2OQL 변환기는 XQL을 입력받아서, 먼저 XQL을 문법에 따라서 파싱을 하고, 파싱할 때 생성되는 파싱 트리를 해석하여 OQL로 변환하게 된다. OQL로 변환할 때는, 파싱 트리의 구조를 해석하여 각 노드가 OQL의 select, from, where 절 중에 어디에 들어갈지를 판단하고, OQL의 select, from, where절을 구성하게 된다.

XQL2OQL 변환기는 <그림 3>에서와 같이 기능이 크게 3부분으로 나뉘게 된다.

- XQL Lexical Analyzer 모듈 : XQL을 입력받아서 토큰을 분리해 낸다.
- XQL Parser 모듈 : 분리된 토큰을 파싱하면서 파싱 트리를 생성한다. 그리고, XQL 문법에 맞는지도 검사한다.
- OQL 생성기 모듈 : 생성된 파싱 트리의 구조를 해석하여 OQL로 변환한다.



<그림 3> XQL2OQL 변환기

3.3 Object2XML 변환기

1) Object2XML 변환기의 개요

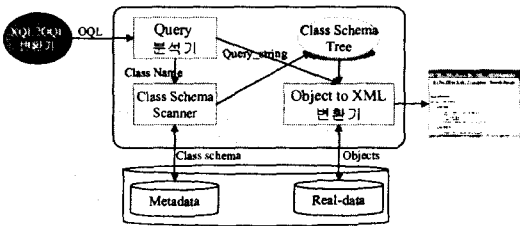
Object2XML 변환기는 객체지향 DBMS에 대한 질의 결과로 얻어지는 객체를 XML형태의 문서로 변환하는 모듈이다. 즉 객체지향 DBMS에 대한 OQL형태의 질의를 입력으로 받아 질의를 수행하고, 결과로써 XML문서를 출력한다. 따라서 변환기 모듈은 질의의 대상이 되는 DBMS의 OQL과 질의 결과 형태에 의존적인 모듈로서 개발되어진다.

2) Object2XML 변환기의 설계

변환기에 대한 주 입력은 XQL2OQL 변환기에 의해 변환된 OQL형태의 질의문이 된다.

DBMS에 대한 질의 결과를 XML로 변환하기 위해서는 정확한 DB 클래스 스키마 정보를 얻을 필요가 있다. 변환기는 이러한 클래스 스키마 정보를 바탕으로 검색된 결과들을 객체 단위로 XML 엘리먼트로 변환한다. 이때 바다-III 스키마 구조와 XML의 엘리먼트 구조는 모두 트리 형태의 구조를 가진다. 따라서 DB2XML에서 사용되었던 별도의 매핑물을 필요로 하지 않는다.

본 논문을 통해 구현된 변환기의 시스템의 구조는 다음 그림과 같다.



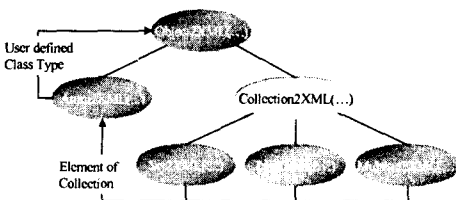
<그림 4> Object2XML 변환기 구조

3) Object2XML 변환기의 구현

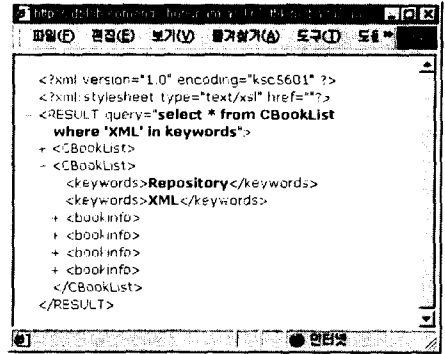
<그림 4>에 나타난 바와 같이 변환기는 다음과 같은 단계를 거친다.

- 단계 1 : 질의 분석
바다-III OQL형태의 질의를 분석하여 변환 대상 클래스의 클래스 이름을 얻는다.
- 단계 2 : 클래스 스키마 정보 획득
질의 분석단계에서 얻은 클래스 이름을 사용하여 DBMS의 Dictionary로부터 클래스 스키마 정보를 얻고, 이를 이용하여 트리 형태의 자료구조로 변환, 유지한다.
- 단계 3 : 질의 결과 변환
입력받은 질의를 실행하여 질의 결과를 객체단위로 변환한다.

앞서 밝힌 바와 같이 바다-III의 스키마구조와 XML의 엘리먼트 구조는 모두 트리 구조를 갖는다. 따라서 변환에 사용되어지는 모듈간에도 이러한 트리 형태의 호출이 일어나며 이러한 트리 형태의 함수호출은 재귀적(recursive)으로 발생한다.



<그림 5> 트리 형태의 변환 함수 호출



<그림 6> IE 5.0 브라우저를 사용한 변환 결과 출력

4. 결론 및 향후 연구 과제

본 논문을 통해 개발된 시스템은 기존의 정보 시스템을 XML 데이터 모델을 적용하여 일관된 방법에 의해 처리하고자 하는 노력에서 출발하였다. 그 중 데이터간의 관계 표현이 용이하며 객체간의 포함관계의 표현이 용이한 객체지향 데이터베이스를 그 대상으로 함으로써 변환된 결과 문서의 활용 가능성을 높일 수 있었다.

차후 바다-III 데이터베이스의 질의 기능의 확장에 따라 지속적인 변환기 시스템의 확장이 필요로 하며 궁극적으로는 데이터베이스 시스템에 독립적인 형태의 변환기 개발이 요구된다. 또한 W3C에 의해 추진되고 있는 XML을 사용한 DBMS 데이터 표현에 관한 표준안인 XML Data, XML Schema, DCD등의 지원을 위한 연구도 계속되어질 것이다[6][7][8][9].

5. 참고문헌

- [1] W3C, Extensible Markup Language(XML) 1.0, REC-xml-19980210, <http://www.w3.org/TR/REC-xml/>, 1998
- [2] C. Baru, A. Gupta, B. Ludascher, R. Marciano, Y. Papakostas, P. Velikhov, V. Chu, "XML-Based Information Mediation with MIX", exhibition program, ACM Conf. on Management of Data (SIGMOD'99), Philadelphia, USA.
- [3] DB2XML home page: <http://www.informatik.fh-wiesbaden.de/~turau/DB2XML/index.html>
- [4] W3C, XML Query Language (XQL), <http://www.w3.org/TandS/QL/QL98/pp/xql.html>, 1998
- [5] 바다-III 객체지향 응용 프로그램 인터페이스 Revision 4.0, 1998. 8. 3, 한국전자통신연구원.
- [6] W3C, XML-Data, <http://www.w3.org/TR/1998/NOTE-XML-data-0105/>, 1998
- [7] W3C, XML Schema Part 1: Structures Working Draft 6-May-1999, <http://www.w3.org/1999/05/06-xmlschema-1/>, 1999
- [8] W3C, XML Schema Part 2: Datatypes Working Draft 06-May-1999, <http://www.w3.org/1999/05/06-xmlschema-2/>, 1999
- [9] W3C, Document Content Description for XML, <http://www.w3.org/TR/1998/NOTE-dcd-19980731>, 1998