

# 링크를 지원하기 위한 XML 질의어의 확장\*

문찬호<sup>0+</sup> 강현철<sup>+</sup> 서상구<sup>++</sup>

<sup>+</sup>중앙대학교 컴퓨터공학과, <sup>++</sup>광운대학교 경영정보학과

<sup>+</sup>{moonch, hckang}@rose.cse.cau.ac.kr, <sup>++</sup>skseo@daisy.gwu.ac.kr

## Extension of a XML Query Language to Support Links

ChanHo Moon<sup>0+</sup> Hyunchul Kang<sup>+</sup> Sang-Koo Seo<sup>++</sup>

<sup>+</sup>Dept. of Computer Science and Engineering, Chung-Ang University

<sup>++</sup> Dept. of MIS, Kwangwoon University

### 요 약

XML 문서의 폭발적인 확산에 따라 Web 상의 각 사이트는 XML 문서를 분산 저장하여 관리하고 있다. 그리고 XML 문서들 간에는 다양한 링크를 통해 관련된 정보를 참조할 수 있다. 그러나 현재까지 XML과 관련하여, 링크를 지원하는 질의어 개발이나 링크를 활용한 XML 검색 시스템 개발에 대한 연구는 미비하다. 본 논문에서는, 실제 Web 상에 분산되어 저장된 전자 문서들 중 하이퍼링크로 연결된 XML 문서들을 대상으로, 링크를 지원하기 위해 기존 XML 질의어를 확장에 대해 연구하였다. 이를 위해 XML 링크를 지원하는 질의어의 유형을 분류하고, 링크를 지원하기 위한 XML 질의어의 확장에 대해 기술한다. 그리고 링크를 활용한 XML 질의 처리 방안에 대해 기술한다.

### 1. 서론

현재의 컴퓨팅 환경에서는 Web의 발전과 더불어 전자 문서의 활용 분야가 점점 확대되고 있다. 특히 인터넷 쇼핑몰, 전자 상거래와 같은 서비스 분야에서는 전자 문서를 통해 다양한 형태의 정보를 제공하고 있다. 이에 따라 전자 문서들에 대한 효과적인 관리 기능의 필요성이 대두되고 있다. 이러한 요구를 충족시키기 위해 XML(eXtensible Markup Language)이 등장하였다[1]. XML은 HTML의 대안으로 W3C에서 제안한 차세대 표준 마크업 언어로서 간단하면서도 다루기 쉽다는 특징을 갖는다.

XML 문서의 폭발적인 확산에 따라 Web 상의 각 사이트는 XML 문서들을 분산 저장하여 관리하고 있다. 그리고 XML 문서들 간에는 다양한 링크를 통해 관련된 정보를 참조할 수 있다.

다양한 정보 형태를 가진 전자 문서의 효과적인 관리를 위해 XML과 관련된 연구들이 현재 활발히 진행되고 있다. XML 문서들을 저장, 관리 및 검색할 수 있는 XML 저장 관리(repository) 시스템 개발에 대한 연구[2], XML과 관련된 질의어에 관한 연구[3][4], 기존 데이터베이스 시스템에 저장된 데이터를 XML 문서로 변환하는 도구 개발에 관한 연구(예를 들어, 관계 데이터베이스에서 XML 문서로, 객체 지향형 데이터베이스에서 XML 문서로)[5] 등이 그 예이다. 이들 XML과 관련된 연구들의 대부분은 XML 문서를 저장하고 있는 XML 저장 관리 시스템에 관한 것이다.

실제 Web 상에는 HTML 문서와 같이 XML 문서들도 서로 하

\* 본 논문은 한국 과학 재단의 특정 기초 연구 지원 사업에 의한 것임.

이퍼링크로 연결되어 있는 상태로 분산되어 있다. 즉, 서로 다른 Web 사이트들 간에도 XML 문서에 대한 링크가 존재한다. 특히 HTML과는 달리 XML에서는 다양한 링크 형식을 제공하고 있다. 본 논문에서는, XML 링크를 지원하는 질의어의 유형에 대해 분류해 보고, 링크를 지원하기 위한 XML 질의어의 확장에 대해 기술한다. 그리고 링크를 활용한 XML 질의 처리 방안에 대해 기술한다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문과 관련된 기존 연구를 살펴본다. 3절과 4절에서는 링크를 지원하기 위한 XML 질의어의 확장과 링크를 활용한 XML 질의 처리 방안을 각각 제시한다. 5절에서 결론을 맺는다.

### 2. 관련 연구

XML 문서는 기존의 정보와는 달리 논리적 구조를 갖는다. 현재 XML과 관련된 질의어는 XQL(XML Query Language)[3], XML-QL[4] 등이 있으며 이들은 XML 문서의 구조적인 특성을 반영한 구조/내용 기반 검색들을 지원하고 있다. 그러나 현재까지 제시된 XML 관련 질의어는 링크를 지원하지 못하고 있다.

XML에서의 링크는 XLink[6]와 Xpointer[7]를 사용하여 각각 문서와 문서 사이의 링크, 문서 내부에서의 연결 정보를 표현하고 있다. HTML에서는 <A> 태그를 이용하여 문서로의 링크를 만들 수 있다. 즉, 한 자원(<A>와 </A>로 둘러싸이는 텍스트 및 이미지)과 URL이 1:1로 링크된다. XML에서도 HTML에서와 같은 링크 형식을 제공하는데 이를 단순 링크(simple link)라 한다((그림 1) 참조). 그리고 XML에서는 한 자원에 대해 여러 개의 URL과 링크할 수 있는 형식을 제공하는데 이

를 확장 링크(extended link)라 한다(그림 2) 참조). 지금까지 XML 문서를 지원하기 위한 브라우저의 개발은 진행되고 있으나, XML 문서를 지원하는 표준화된 브라우저는 현재 존재하지 않는다. 이에 따라 XML에서의 확장 링크를 브라우저에서 어떻게 처리할 지는 알 수 없지만, HTML에서의 단순 링크의 제약을 확장 링크를 사용하면 좀더 정교하게 링크를 표현할 수 있을 것이다. 현재까지 XML과 관련하여 링크를 지원하는 질의어 개발이나 링크를 지원하기 위한 검색 시스템 개발에 대한 연구는 미비하다.

```
<xdoc>
<para>
This is the report of on the origins of the cold war.
<locator xml:link="simple" href="intro.xml">Introduction</locator>
<locator xml:link="simple" href="propos.xml">Proposal</locator>
<locator xml:link="simple" href="summary.xml">Summary</locator>
</para>
</xdoc>
```

(그림 1) XML에서의 단순 링크 예

```
<xdoc>
<para>
This is the report of
<mylink xml:link="extended" inline="true" title="Debate">
<locator xml:link="locator" href="intro.xml" title="Introduction"/>
<locator xml:link="locator" href="propos.xml" title="Proposal" />
<locator xml:link="locator" href="rebut.xml" title="Rebutal" />
a debate
</mylink>
on the origins of the cold war.
</para>
</xdoc>
```

(그림 2) XML에서의 확장 링크 예

### 3. 링크를 지원하기 위한 XML 질의어의 확장

#### 3.1 XML 링크를 지원하는 질의의 유형

Web 상에 두 XML 문서가 링크로 연결된 경우, XML 문서는 각각 링크로 상대 문서를 참조하는 문서와 링크로 참조되는 문서로 구분할 수 있다. 링크로 상대 문서를 참조하는 XML 문서를 대상으로, 사용자 인터페이스를 통해 입력받은 질의를 수행하는 사이트를 질의 제기 사이트라 하면, 링크로 참조되는 XML 문서는 질의 제기 사이트 내에 존재할 수도 있고 다른 사이트에 존재할 수 있다. 이에 따라 링크를 지원하는 질의는 다음과 같이 크게 세 가지 유형으로 분류할 수 있다.

(유형 1) 링크로 참조되는 XML 문서가 질의 제기 사이트 내에 있는 경우

A 쇼핑몰은 판매되는 총 물품에 대한 카탈로그 정보를 갖는 XML 문서와 각 물품에 대한 상세한 정보를 갖는 XML 문서들을 관리한다고 하면, (유형 1)의 예는 다음과 같다.

```
A 쇼핑몰 사이트 내 카탈로그 XML 문서를 대상으로, X
물품에 대한 상세한 정보를 찾으시오.
```

(유형 2) 링크로 참조되는 XML 문서가 질의 제기 사이트가 아닌 사이트에 있는 경우

한국 학회는 회원으로 등록된 회원들에 대한 정보를 갖는 XML 문서를 관리하고, 회원들의 홈페이지는 한국 학회가 관리하는 사이트가 아닌 다른 사이트에 존재한다고 하면, (유형 2)의 예는 다음과 같다.

```
한국 학회 사이트 내 회원 정보 XML 문서를 대상으로,
회원 중 이름이 홍길동이라는 사람의 홈페이지를 찾아
bibliography 정보를 찾으시오.
```

(유형 3) 링크로 참조되는 XML 문서가 질의 제기 사이트와 다른 사이트에 모두 존재할 경우

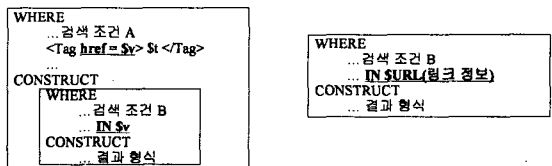
물품에 대한 가격 비교 정보를 제공하기 위해, A 쇼핑몰에서는 타 쇼핑몰이 판매하는 동일 물품에 대한 정보를 제공한다고 하면, (유형 3)의 예는 다음과 같다.

```
A 쇼핑몰 사이트 내 카탈로그 XML 문서를 대상으로, X
물품에 대한 상세한 정보와 타 쇼핑몰에서 판매되는 X
물품에 대한 상세한 정보를 찾으시오.
```

#### 3.2 링크를 지원하기 위한 XML 질의어의 확장

본 절에서는 링크를 지원하기 위한 XML 질의어의 확장에 대해서 기술한다. 이를 위해 XML과 관련된 질의어들 중 서로 다른 사이트 간의 XML 문서에 대한 연산을 쉽게 표현할 수 있는 XML-QL[4]을 확장하여, 링크를 지원하는 검색 질의를 표현할 수 있도록 한다.

(그림 3) (a)는 링크를 지원하기 위한 XML 질의어의 구문(syntax)을 나타낸 것이다. 링크를 지원하기 위한 XML 질의어의 구문은 크게 WHERE 절과 CONSTRUCT 절로 구성되며, CONSTRUCT 절 내에 WHERE 절과 CONSTRUCT 절이 중첩되는 구조를 갖는다. WHERE 절은 SQL의 SELECT-WHERE 구조에서의 WHERE절에 해당하고, XML 문서를 대상으로 한 검색 조건을 나타낸다. CONSTRUCT 절은 질의 수행 결과를 XML 문서로 나타내기 위한 결과 형식을 나타낸다. 질의어 내의 "href=\$v"와 "IN \$v" 구문은 질의가 링크를 지원하는 검색 질의임을 나타낸다. \$v는 질의 수행 시 검색 대상이 되는 URL을 저장하는 변수이다. "href=\$v"는 애트리뷰트(attribute)로 링크를 포함하고 있는 XML 문서의 엘리먼트(element)를 찾아 변수 \$v에 해당 URL을 저장하라는 의미를 갖는다. "IN \$v"는 변수 \$v가 갖는 URL의 XML 문서를 대상으로 검색을 수행하라는 의미를 갖는다. (그림 3) (a)에 기술한 질의어 구문은 <검색 조건 A>에 부합되는 엘리먼트의 링크를 활용하여 해당 링크의 XML 문서를 찾고 <검색 조건 B>에 부합되는 내용을 찾아 결과를 리턴하라는 검색 요청을 표현하고 있다.



(a) 링크를 지원하기 위한 XML 질의어의 구문

(b) 부질의의 템플릿

(그림 3) 링크를 지원하기 위한 XML 질의어의 구문 및 부질의의 템플릿의 형태

링크를 지원하는 XML 질의의 최종 목표는, 기존 검색 엔진에서 키워드 검색의 결과로 제공하는 링크의 URL을 찾는 것이 아니라 링크로 연결된 상대 XML 문서 내에서 검색 조건에 부합되는 내용을 추출하여 새로운 결과 XML 문서로 리턴하는 것이다. 링크를 지원하기 위한 XML 질의를 수행한 결과는 (그림 3) (a)에서 나타난 <결과 형식>의 XML 문서가 된다. 이를 위해, 질의 제기 사이트는 사용자 인터페이스를 통해 받아들인 질의를, 링크를 지원하기 위한 질의만을 처리하도록 수정한 질의로 수정하여 사용한다. 이를 부질의라 한다. 부질의는 관련된 링크 수만큼 생성된다. 각 부질의는 해당 링크의 URL이 포함된 사이트로 전달되어 해당 사이트에서 질의를 수행하고, 그 결과는 XML 문서로 질의 제기 사이트에게 리턴된다. 부질의를 생성하는 과정은 다음과 같다. 먼저, 사

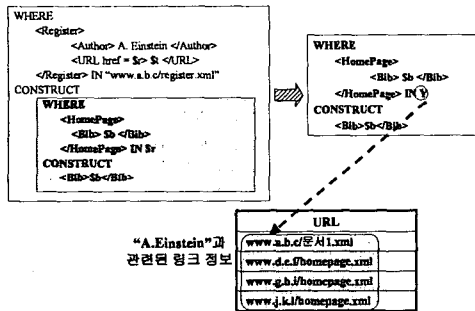
용자 인터페이스를 통해 입력받은 질의로부터 부질의 템플릿을 만든다. (그림 3) (b)는 (그림 3) (a)의 질의에서 부합된 링크들을 찾아 각 링크에 해당하는 부질의를 생성할 수 있도록 만든 부질의 템플릿이다. 부질의 템플릿은 (그림 3) (a)의 음영 표시 부분에서 쉽게 유도할 수 있다. 부질의는 생성된 부질의 템플릿에 링크를 추가함으로써 자동으로 생성할 수 있다.

**4. 링크를 활용한 XML 질의 처리 방안**

본 절에서는 링크를 지원하기 위한 XML 질의의 예를 이용하여, XML 질의 처리 방안에 대해 기술한다. 링크를 지원하기 위한 XML 질의의 예는 다음과 같다.

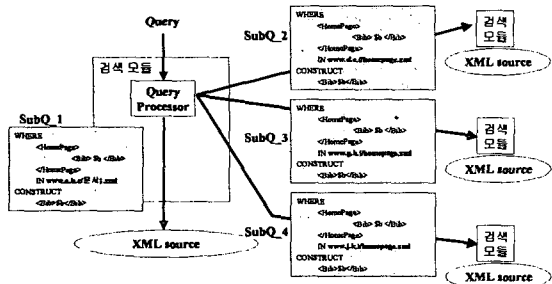
www.a.b.c/register.xml 내 등록된 사람 중 이름이 "A. Einstein" 인 사람의 홈페이지를 찾아 bibliography 정보를 찾으시오.

이와 같은 질의는 Web을 대상으로 빈번히 발생할 수 있다. 위의 질의가 사용자 인터페이스를 통해 질의 제기 사이트로 전달되면, 질의 제기 사이트는 해당 질의로부터 부질의 템플릿을 만들고, 검색 질의에 부합되는 링크를 찾는다. (그림 4)는 링크를 지원하기 위한 질의 및 부질의 템플릿, 그리고 질의에 부합된 링크의 URL을 테이블로 나타낸 것이다.



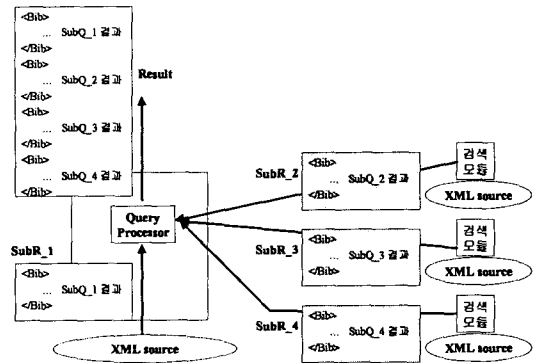
(그림 4) 링크를 지원하기 위한 질의 및 부질의 템플릿 예

부질의 템플릿에 링크를 추가함으로써 부질의를 생성한다. (그림 5)는 부질의의 생성 및 전달 과정을 나타낸 것이다. (그림 5)에 나타난 것과 같이 생성되는 총 부질의의 개수는 4개이다. 생성된 부질의들은 해당 링크의 URL이 포함된 사이트로 전달된다. (그림 5)에 나타난 것과 같이 각 사이트는 XML 문서들을 저장하고 있는 XML 소스(source)와 질의를 처리할 수 있는 검색 모듈로 구성된다. [8] 참조). 검색 모듈의 질의 처리기(Query Processor)는 부질의의 생성 및 전달을 수행한다.



(그림 5) 부질의의 생성 및 전달 과정

부질의를 전달받은 해당 사이트의 검색 모듈은 부질의를 처리하여 그 결과를 XML 문서로 작성하여 질의 제기 사이트로 리턴한다. 마지막으로 질의 제기 사이트는 각 부질의의 결과들을 하나의 결과 XML 문서로 작성하여 사용자에게 리턴한다. (그림 6)은 부질의에 대한 수행 결과 후 사용자에게 최종 XML 문서로 리턴되는 과정을 나타낸 것이다.



(그림 6) 부질의의 수행 결과 후 결과 리턴 과정

**5. 결론**

본 논문에서는, 실제 Web 상에 분산되어 저장된 전자 문서들 중 하이퍼링크로 연결된 XML 문서들을 대상으로 사용자 질의에 대해 효율적인 검색을 지원해주기 위해, 링크를 지원하기 위한 질의어의 확장을 제시하였다. 질의 처리 과정 중 링크와 관련된 질의는 해당 링크를 이용한 부질의를 생성하여 처리한다. 링크와 관련된 XML 문서 검색을 지원하기 위해 본 논문에서는, XML 링크를 지원하는 질의의 유형에 대해 분류해 보았고, 링크를 지원하기 위한 XML 질의어의 확장 및 부질의의 생성에 대해 기술하였다. 그리고 링크를 활용한 XML 질의의 처리 방안에 대해 기술하였다.

**6. 참고문헌**

- [1] T. Bray et al., "Extensible Markup Language (XML) 1.0," <http://www.w3.org/TR/1998/REC-xml-19980210>, 1998.
- [2] C. Baru et al., "XML-Based Information Mediation with MIX," Proc. of the 1999 ACM SIGMOD Int'l Conf. on Management of Data, 1999, Jun., pp. 597-599.
- [3] J. Robie et al., "XML Query Language (XQL)," <http://www.w3.org/TandS/QL/QL98/pp/xql.html>, 1998.
- [4] A. Deutsch et al., "XML-QL: A Query Language for XML," In Proc. of the Query Language Workshop (QL98), 1998, Dec., <http://www.w3.org/TR/1998/NOTE-xml-ql-19980819>.
- [5] D. Florescu and D. Kossmann, "Storing and Querying XML Data using an RDBMS," Bulletin of the Technical Committee on Data Engineering, Vol. 22, No. 3, 1999, pp. 27-34.
- [6] S. DeRose et al., "XML Linking Language(Xlink)," <http://www.w3.org/TR/xlink>, 1999.
- [7] S. DeRose et al., "XML Pointer Language (Xpointer)," <http://www.w3.org/TR/WD-xptr>, 1999.
- [8] 문찬호, 강원철, "링크 정보를 활용한 XML 문서의 검색," 한국정보처리학회, 2000년 춘계 학술발표논문집, 2000.