XML 시각화를 통한 자동 질의문 생성

송민영* 이기호
이화여자대학교 컴퓨터학과
(navyblue_khlee@ewha.ac.kr)

Auto generation of a query language through visualization of XML

Minyoung Song* Kiho Lee
Dept. of Computer Science & Engineering, Ewha Womans University

요약

이를 위해 본 논문에서는 다음과 같이 세 가지 컴퓨터트 기반으로 자동 질의 생성을 위한 visual query interface를 제안하였다. 이 컴퓨터트는 개층적 XML 문서 구조를 보유하는 XML structure viewer로서 사용되는 연산자동화 아이콘화한 operator panel, 그리고 이 두 컴퓨터트의 상호작용으로 작성된 질의 구문을 블록별로 검색하기 위한 query statements panel을 구성한다. 이를 통해 다양한 계층적인 XML 문서의 구조와 복잡한 XML 질의 언어의 구문에 대해서 잘 모르는 초보자도 쉽게 XML 문서 상에서 질의문을 할 수 있다.

1. 서론

XML 관련 신기술 데이터를 표현하는 기초 언어로 XML을 이용한 문서관리 시스템이 확산되고 있다. IT 분야뿐만 아니라 인터넷 관련 기업들을에서도 XML 퍼져난의 활용이 증가하고 있는 한편 다양한 신기술의 등장에 따라 XML 문서에서 질의를 하고자하는 사용자의 요구가 증가하고 있다. 이에 따라 XML 데이터를 대상으로 정보를 추출하고 변환하고 통합하기 위한 질의 언어의 필요성은 대단히 높고, 다양한 XML 퍼져난 언어의 복잡함이 대단하게 되었다. 이러한 XML 문서의 선택 방법은 질의 언어의 구조를 이해하고 사용법을 숙지해야하는 어려움이 있어 일반 사용자에게는 적합하지 않다.

또한 텍스트 형태의 XML 문서는 계층적 구조로 일반 사용자들이 인식하는데 어려움이 있다. 이와 같이 XML 문서 상에서 질의를 할 때, 사용자의 인지적 개념과 질의 언어의 조작 사이에는 큰 차이가 있다. 즉, 사용자는 XML 텍스트에 대한 정확한 이해가 필요하며, 텍스트의 기반의 질의 언어의 복잡한 문법을 숙지해야한다. 따라서 XML 질의 언어를 사용함으로서 사용자의 인지적 개념을 그대로 모형화 할 수 있도록 하고, 각 요소를 사용하기 쉬운 XML 기반에 준 통합 환경을 제공하여 다양한 사용자들에게 기반적인 자식만으로도 쉽게 질의문을 작성 및 생성할 수 있는 시스템 개발이 필요하다.

본 연구는 다른 visual query language [1, 2, 3, 4]와는 달리 XML을 기반으로 하고 있으며, XML 문서의 체계화된 구조를 보여주기 위하여 XML 문서 본문을 통한 structure viewer를 제공함으로써 XML 문서에 대한 구조를 잘 모르는 사용자가 하더라도 쉽게 반영과도 간단한 표현을 제공한다. 또한 이러한 XML 질의 언어의 필요성을 인식한 사용자들이 이용하기 위한 operator panel과 각 expression을 expression panel을 제작함으로써 XML 질의 언어의 구문에 대해서 잘 모르는 초보자라도 쉽게 질의문을 할 수 있도록 한다. 그리고 이러한 각각의 패널들을 잘 조합해 사용하여 질의 언어로 변환하고, 완성된 질의문을 블록별로 응바른 결과를 검색하기 위한 query statements panel을 제공한다. 따라서 사용자는 시각환경에 서의 단순 조작만으로 원하는 결과해석을 생성할 수 있다.

2. 관련연구

XMLExtensible Markup Language)이 자세히 웹 문서의 표준을 만드는 언어로 정해진 이후. XML과 관련하여 XML 문서로 사용자들이 원하는 정보를 얻기 위한 XML 질의 언어에 대한 연구가 활발히 수행되고 있다.

현재 XML과 관련된 질의 언어는 XML-QL, XQL, XQuery 등이 표준화 과정중에 있으며 이들은 XML 문서의 구조적인 특성을 반영한 구조/내용 기반 접근을 지원하고 있다.

또한, 기존의 다양한 그래픽 질의 언어에 대한 연구가 있다. 대표적인 것은 VISUAL, VODDOLE, PESTO [1,2,3,4]등이 있으며 이들은 XML을 대상으로 질의문을 특성화하였다.
사용자 인터페이스에 입력이 들어오면 각 프레임간 상호 작용을 투명하게 입력된 구문을 XQuery로 변환하고 문법에 맞게 정정한다.

3.1 XML structure viewer
XML 문서의 형식은 문서 내용을 '문서 형식 정의' (Document Type Definition, DTD)과 '문서 구성'으로 나눌 수 있다. 문서 형식 정의는 '엘리먼트'를 정의하고 '속성'을 구성하는 '구조'를 정의하며, 현실적인 구조에서 사용하는 XML 문서는 이 두 가지 요소를 공유하게 된다. DTD는 문자 내용과 요소 정의의 구조를 정의할 때에 문서 구조에 만족하는 부분만을 정의하며, XML 문서는 XML 구조화된 구조를 표현하기 위해 문서 내에 있는 요소들의 구조 이름을 XML 문서의 접두사 및 구조의 부분을 사용하여 단위별로 표현될 수 있다.

본 시스템은 입력 문서의 DTD 를 통한 문서의 채택 과정을 거쳐서, DTD의 접두사와 목록을 사용하는 XML 문서 구조를 제공할 때에 문서 구조에 만족하는 부분만을 사용하여 수학적 구조를 표현한다. DTD는 단일 요소로 구성된 XML 문서의 구조를 표현하고, 문서의 구조를 표현하기 위해, XML 문서의 구조를 XML 문서의 접두사 및 구조의 부분을 사용하여 단위별로 표현할 수 있다.

3.2 Visualized operator panel
본 시스템은 XQuery에서 사용하는 operator와 FLWR 구문을 표현할 수 있도록 expression panel과 대본문의 형식을 제공하는 산술, 비교, 논리, 배열, 조건 연산자를 가진 '함수'를 제공한다. 산술 연산을 위해 'add', 'div', 'mod 연산자를 포함하고 있고, 비교 연산자에는 값 비교, 논리 비교로 구성되어 있다. 연산자로 다음과 같이 구현되어 있다: eq, ne, lt, le, gt, ge, ==, !=, <, >, BEFORE, AFTER가 있다. 논리 연산을 위해서 `and`, `or` 연산자를 제공하며, 배열 연산은 'for', 'union', 'intersect', `except` 연산자를 제공한다. sorted 연산자와 조건을 제공하기 위한 if-then-else를 지원한다.
단 함수로 avg, sum, count, max, min, contains를 제공한다. 그림 3은 operator 패널의 설계화면이고 그림 4는 expression 패널의 설계화면이다. 각각은 사용자 변수와 조건을 입력받기 위한 텍스트 필드와 버튼으로 구성되어 있다.

경별을 하기 위한 sorting expression 패널에서는 descending, ascending 선택버튼을 사용하여 오름차순 혹은 내림차순 정렬을 가능하게 하고, 조건문을 만들기 위한 conditional expression 패널은 if-then-else 구조를 지원한다. 또한 XQuery에서는 SQL의 SELECT-FROM-WHERE 절과 비슷한 FOR-LET-WHERE-RETURN 절을 사용하는데 이를 처리할 수 있도록 FLWR expression 패널에는 내부의 패널을 포함하고 있다. quantified expression 패널은 predicate를 만드는 노드를 검색하기 위한 것으로 조건을 만족하는 모든 노드를 검색하기 위한 EVERY 선택 버튼과 적어도 하나의 노드를 검색하기 위한 SOME 선택 버튼으로 구성된다. 그리고 조건을 명시해주기 위한 in-satisfies 구문을 제공한다. typeswitch expression 패널은 표준 데이터타입뿐 아니라 사용자 정의 데이터타입인 XML 스키마의 모든 데이터 타입을 지원하기 위해 typeswitch-as-default-return 구문을 제공한다.

3.3 Query statements panel
Query statements 패널은 각 연산들간의 상호작용을 통해 입력된 값을 결과로 변환하고 전달된 구문이 컴퓨터의 이해가 가능한 형태로 표현된 operator 패널과 FLWR expression에 있는 버튼에 의해 선택되고, 이들로 연산하는 operator는 operator 패널에서 사용하여 이들로 조합한 결과가 바로 이 Query statements 패널에 텍스트로 표시된다.

4. 프로토타입 구현 및 실행
전체 화면 구성은 배뉴버와 네 부분의 프레임들로 이루어져 있다. 그림 5는 slideSample.xml 문서를 대상으로 slide-title의 item 엘리먼트에 "XQuery" 비주얼을 포함하고 있는 slide-title을 오름차순으로 출력하라는 XQuery 질의문을 생성하는 예시이다.

5. 결론 및 향후 연구
인터넷 문서의 표준 사양인 XML 문서가 늘어나면서 XML 문서를 효과적으로 관리하고 검색하기 위한 시스템이 개발되고 있지만 지금까지 개발된 XML 문서의 검색 시스템은 절의 구조를 이해하지 못한 일반 사용자가 사용하기에는 어려움이 있다. 따라서 사용자가 쉽게 사용할 수 있으며 다양한 결과를 가진 시스템이 요구되었고 이에 대한 요구사항에 의해 본 논문에서는 사용자가 보다 쉽게 XML 문서상에서 질의를 할 수 있도록 하고자 자동 생성을 위한 visual query interface를 제안하였다.

향후 연구 과제로는 본 설계를 바탕으로 시스템을 구현하고, XML 프로세서와의 연동을 통한 결과를 시각화하는 부분에 대한 연구가 필요하다. 또한 이를 기반으로 다음 XML의 연어에도 응용이 가능하다고 생각된다.