

# GML 문서의 통합 지리 정보 검색을 위한 XQuery의 확장

안영수<sup>o</sup> 박순영 정원일 배해영  
인하대학교 전자계산공학과

{owlet76<sup>o</sup>, sunny. wncung}@dblab.inha.ac.kr, hybae@inha.ac.kr

## A Extension of XQuery for Integrated Geographic Information Retrieval of GML

Young-Su An<sup>o</sup>, Soon-Young Park, Warnill Chung, Hae-Young Bae  
Dept. of Computer Science & Engineering, Inha University

### 요 약

GML(Geography Markup Language)은 XML(eXtensible Markup Language)의 장점을 이용하여 지리 데이터의 저장과 전송을 위해 OGC(Open GIS Consortium)에서 제안되었다. 이러한 GML은 웹 환경의 발전에 따라 지리 정보 통합을 위해 널리 이용되고 있으나, GML 문서에 대한 지리 데이터 검색 기술은 미흡하다. 최근 W3C(World Wide Web Consortium)에서는 XML 문서로부터 데이터를 추출하기 위한 표준 질의어로 XQuery를 추천하기 위한 작업 초안(working draft) 상태에 있다. 그러나 이러한 XML 질의어는 지리 정보의 추출에 대한 고려는 하고 있지 않다.

따라서 본 논문에서는 지리 정보를 포함하고 있는 GML 문서를 기반으로 지리 정보를 검색, 추출할 수 있는 GML Query language인 GQuery를 제안하고자 한다. 본 논문의 GQuery는 서로 이질적인 공간 데이터베이스의 데이터를 통합하여, 통합된 문서내의 지리정보를 추출하기 위한 사용자 인터페이스로 이용할 수 있다.

### 1. 서 론

GML[1]은 XML[3]의 한 부분으로 지리 정보를 표준화된 형식으로 표현, 응용할 수 있도록 OGC에 의해 제안된 표준이다. GML은 웹 환경에서의 지리 데이터 통합과 문서 교환의 표준 데이터 형식으로 이용되고 있다. 이에 따라 수많은 GML 문서의 지리 데이터를 효율적으로 추출할 수 있는 질의어에 대한 연구가 필요하게 되었다. 현재 XML 문서를 검색하기 위한 대표적인 질의어로는 Lorel, XML-QL, XML-GL, XSL, XQL, XQuery 등이 소개되고 있다[4, 6]. 이들 중 XQuery는 W3C에 의해 표준화가 진행중인 XML 질의어로 Quilt[7]를 기반으로 하고 있다. Quilt는 XML-QL, XPath[5], XQL 및 몇몇 다른 XML 질의어의 개념과 SQL, OQL의 개념을 추가하여 확장한 언어이다. 이러한 Quilt를 기반으로 하는 XQuery는 강력한 질의어로 최근 XML 질의어의 표준으로 간주되고 있다. 하지만 이러한 XML 질의어 들은 지리 정보에 대한 고려를 하고 있지 않다. 따라서 지리정보와 비 지리정보를 포함하고 있는 GML 문서의 특성을 고려한 데이터 검색을 위한 질의어가 요구 된다.

본 논문<sup>1</sup>에서는 GML 문서의 지리 정보 검색을 위하여 여러 XML 질의어들 중 다양한 조건 지원과 복잡한 질의에 유용한 포괄적 언어로 최근 표준으로 간주되고 있는 XQuery를 공간 SQL을 지원하

도록 확장하여 GQuery를 제안 한다.

본 논문의 구성은 다음과 같다. 제 2장 관련 연구에서는 본 연구의 기본이 되는 XQuery에 대해 살펴보고, 제 3장에서는 본 논문에서 제안하는 GQuery의 공간 SQL과 GQuery 구조에 대해 예제와 함께 살펴보도록 한다. 마지막으로 제 4장에서는 결론 및 향후 과제를 논한다.

### 2. 관련 연구

XML 질의어는 데이터베이스와 같이 XML 문서나 XML 문서의 컬렉션에 접근할 수 있다. W3C XML Query 워킹 그룹에서는 웹에서 XML 문서로부터 데이터를 추출하기 위한 유연한 질의 기능을 제공하기 위한 목적으로 XQuery 명세를 제공한다. XQuery는 표현(expression)들로 구성되어 있으며, 이들 중 가장 주목할 만한 경로 표현과 FLWR 표현에 대해 살펴 본다.

#### 2.1 경로 표현(path expression)

XQuery 경로 표현은 XPath의 축약형 구문을 사용하여 효과적으로 지역 경로 내에서 사용하는 축의 수를 줄일 수 있도록 한다. 경로 표현은 질의로서 출로 사용될 수 있으며, 사용자들이 쉽게 질의를 읽고 쓸 수 있어야 한다는 요구 사항을 만족하며 URL 상에 쉽게 삽입 될 수 있는 장점을 가지고 있다. [그림 1]은 'authors.xml' 문서에서 2001년에 출판된 모든 도서를 가져오라'는 질의문을 경로 표현으로 나타낸 그림이다.

<sup>1</sup> 본 연구는 대학 IT연구센터 육성 지원사업의 연구 결과로 수행 되었음.

```
document(" authors.xml" )/authors/author/book[year=" 2001" ]
```

[그림 1] 경로 표현 형태의 XQuery

2.2 FLWR 표현(For Let Where Return)

XQuery의 FLWR 표현은 경로 표현의 취약점을 보완하여 노드 컬렉션의 반복, 변환, 임의의 결과 구성, IF-THEN-ELSE의 조건적인 처리, Inner/Outer 조인, 질의문 내에 함수, 사용자 정의 함수를 통한 확장 등을 제공한다. 또한 수학적, 논리적인 기본 연산자와 비교 연산자, UNION/INTERSECT/EXCEPT 연산자 등을 제공한다. XQuery는 SQL 사용자에게 친숙하도록 SQL과 같은 구문으로 설계되었다. [그림 2]는 '출판된 도서 연도를 포함하고, 최근 출판된 도서 연도보다 10년 이내의 것만 포함하는 새로운 well-formed XML 문서를 생성하라'는 질의문을 FLWR 표현으로 나타낸 그림이다.

```
FOR $y IN distinct (/authors/author/*/year)
LET $d := max(/authors/author/*/year) - $y
WHERE $d < 10
RETURN <publishYear> $y/text() </publishYear>
```

[그림 2] FLWR 표현 형태의 XQuery

3. 지리 정보 검색을 위한 XQuery의 확장

본 논문에서 XQuery 확장을 위해 적용한 공간 SQL은 GMS<sup>2</sup>[8]에서 제공하는 공간 SQL을 기반으로 한다. GMS는 기존의 SQL-92를 기반으로, 공간 데이터 처리를 위해 확장한 공간 SQL을 지원한다. GMS의 공간 SQL은 OGC에서 표준으로 제안하는 공간 데이터 타입 및 공간 관계 연산자, 공간 함수를 지원한다[2].

3.1 GMS의 공간 SQL

GMS의 공간 SQL의 객체 모델은 POINT, LINESTRING, POLYGON과 이들간의 조합인 MULTIPOINT, MULTILINESTRING, MULTIPOLYGON, GEOMETRYCOLLECTION 타입으로 구성 된다.

● 공간 연산자: 공간 관계 연산자는 두 개의 공간 객체를 대상으로 객체 간의 위상 관계를 연산하여 그 결과를 반환 한다. 객체 위상 관계는 [표 1]과 같으며 결과는 TRUE/FALSE 이다.

[표 1] 기본적인 공간 연산

공간 연산자	연산
EQUALS	두 객체 타입이 같고 서로 좌표가 같으면 TRUE
DISJOINT	두 객체 사이의 교집합이 없으면 TRUE
INTERSECTS	두 객체 사이의 교집합 존재 시 TRUE
TOUCHES	두 객체의 interior 영역에 대해 어떠한 공통된 좌표가 존재하지 않고 하나의 boundary 영역과 다른 하나의 interior나 boundary 영역에 교집합이 존재하면 TRUE
OVERLAPS	객체의 타입이 같고 두 객체의 교집합이 두 객체와 다르지만, 결과가 같은 타입의 객체인 것이 존재하면 TRUE
CROSSES	교집합의 결과가 두 객체의 최대 차원보다 작은 차원을 가지고, 두 객체의 interior 영역에 존재하면 TRUE

WITHIN	첫 번째 객체가 완전히 두 번째 객체 안에 존재하면 TRUE
CONTAINS	두 번째 객체가 완전히 첫 번째 객체의 영역 안에 포함되면 TRUE

이러한 공간 연산자 외에 GMS에서는 사용자에게 빠른 응답 시간을 주기 위한 [표 2]와 같은 공간 연산자를 제공한다.

[표 2] 추가적인 공간 연산

ENVELOPEDINTERSECT	공간 객체 MBR 필더링 단계에서 후보 객체로 선정된 객체만을 검색하기 위한 연산(Index가 없는 경우)
INDEXINTERSECT	공간 객체 MBR 필더링 단계에서 후보 객체로 선정된 객체만을 검색하기 위한 연산

● 공간 함수: 공간 함수는 데이터 항목을 조작하고 결과를 반환한다. 이는 점에서 연산자와 유사하나 내부적으로 복잡한 단계의 연산을 거쳐 결과를 반환하므로 단순 연산자로 표현할 수 없는 결과를 얻어낼 경우 유용하다. 공간 함수는 단일 객체 함수, 거리관계 함수, 객체 생성 함수로 세분화 할 수 있다.

[표 3] 단일 객체 함수

객체 함수	
Dimension	공간 객체의 차원을 반환
GeometryType	공간 객체의 타입을 문자열로 반환
AsText	공간 객체의 내부 표현을 표준 텍스트 표현으로 변환하여 반환
AsBinary	공간 객체의 내부 표현을 표준 바이너리 표현으로 변환하여 반환
IsEmpty	공간 객체가 Empty Set이면 TRUE
IsSimple	공간 객체가 Simple한 특성을 가지고 있으면 TRUE
Boundary	공간 객체의 Combinational Boundary를 반환
Envelop	공간 객체를 둘러 싸고 있는 최소 경계 사각형 반환
POINT 함수	
X	Point 객체의 X 좌표를 반환
Y	Point 객체의 Y 좌표를 반환
LINESTRING 함수	
StartPoint	Linestring 객체의 시작점을 반환
EndPoint	Linestring 객체의 끝점을 반환
IsClosed	Linestring 객체의 시작점과 끝점이 같으면 TRUE
IsRing	Linestring 이 Ring 형태를 가진 객체이면 TRUE (조건은 객체가 Close 되고, Simple 한 경우)
Length	Linestring 객체의 길이를 반환
NumPoints	Linestring 객체의 점의 개수를 반환
PointN	Linestring 객체에서 n 번째 point 반환
POLYGON 함수	
Centroid	Polygon 객체의 중심점을 구해서 반환
Area	Polygon 객체의 면적을 구해서 반환
ExteriorRing	Polygon 객체의 바깥쪽 Ring을 반환
NumInteriorRing	Polygon 객체의 Interior Ring의 개수를 반환
InterRingN	Polygon 객체에서 n번째 Interior Ring을 반환
GeometryCollection 함수	
NumGeometries	Collection 객체에서 Geometry의 개수를 반환
GeometryN	Collection 객체에서 n 번째 Geometry 객체를 반환
MULTILINESTRING 함수	
IsClosed	Multilinestring 객체가 close 되어있으면 TRUE
Length	Multilinestring 객체의 전체 길이를 반환
MULTIPOLYGON	
Centroid	Multipolygon 객체의 중심점을 반환
Area	Multipolygon 객체의 면적을 반환

[표 4] 거리 관계 함수

Distance	geometry1 객체와 geometry2 객체 사이의 거리를 반환
----------	---------------------------------------

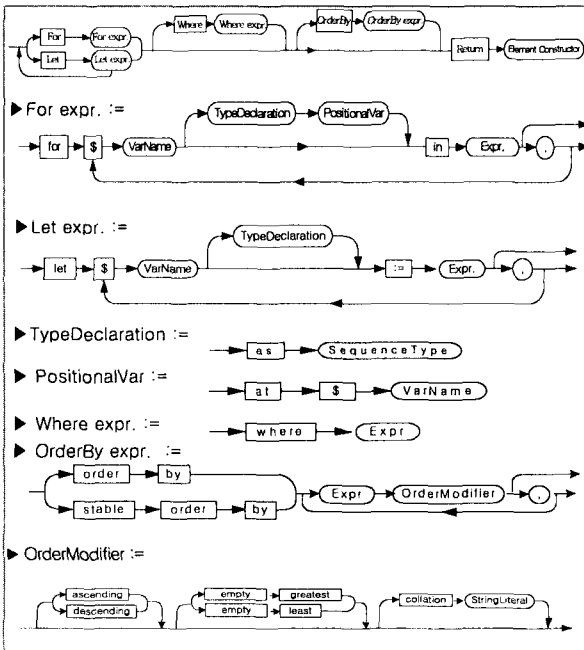
<sup>2</sup> (주)지오매니아에 의해 상용화된 공간 DBMS.

[표 5] 객체 생성 함수

Intersection	geometry1 객체와 geometry2 객체의 교집합 영역 반환
Difference	geometry1 객체와 geometry2 객체의 차집합 영역 반환
Union	geometry1 객체와 geometry2 객체의 합집합 영역 반환
SymDifference	geometry1 객체와 geometry2 객체의 대칭 차집합 영역 반환
Buffer	geometry 객체를 d의 크기만큼 Buffering
ConvexHull	geometry 객체의 Convex Hull 영역을 반환

3.2 GML Query language(GQuery)

GQuery의 기본적인 구문은 XQuery의 구문을 따른다. EQUALS, DISJOINT, INTERSECTS, TOUCHES 등과 같은 두 공간 객체를 대상으로 객체간 위상 관계를 연산하는 공간 연산자는 WHERE 절에서 사용하며, 단일 객체 함수, 거리함수, 객체 생성 함수는 FOR, LET, WHERE, RETURN 절에 사용한다. 이러한 GQuery는 GML 문서에 대한 검색을 제공하며, 추출한 결과를 RETURN 절의 엘리먼트 생성자를 이용하여 새로운 문서로 구성할 수 있다. [그림 3]은 GQuery의 기본적인 구문 구성이다.



[그림 3] GQuery 기본 구조

다음은 [그림 3]의 구문을 바탕으로 공간 질의를 위한 GQuery의 예제이다. GMS의 공간 SQL 형태를 GQuery 형태로 변환한다.

[예제 1] Polygon 객체의 면적을 구해서 반환

```
SELECT AREA(obj) FROM 건물 WHERE 건물명 = '한국은행';
FOR $a IN document("건물.xml")//building[name="한국은행"]
RETURN <KoreaBank>
    <area>AREA($a/obj)</area>
</KoreaBank>
```

[예제 2] 강 5000km 이내에 위치하는 도시의 이름

```
SELECT CityName
FROM 도시, 강
WHERE Overlap(도시.obj, Buffer(강.obj, 5000))

FOR $a IN document("도시.xml")//city,
    $b IN document("강.xml")//river
WHERE Overlap($a/obj, Buffer($b/obj, 5000)) eq true
RETURN
    <Within5000CityName>$a/name</Within5000CityName>
```

[예제3] multilinestring 객체의 전체 길이 반환

```
SELECT SLENGTH(obj) FROM 도로;
LET $a := document("도로.xml")//road
RETURN <totalLength>SLENGTH($a/obj)</totalLength>
```

이와 같이 제안한 GQuery는 분산환경의 서로 이질적인 하부 공간 데이터베이스의 공간 데이터를 통합하여, 통합된 문서내의 지리정보를 추출하기 위한 사용자 인터페이스로 이용된다. 따라서 사용자는 이러한 장점을 이용하여 이질적인 공간 데이터베이스를 고려하지 않고 GQuery를 사용하여 원하는 데이터를 추출할 수 있다.

4. 결론 및 향후 과제

본 논문에서는 GML 문서 내의 지리 데이터 추출을 위해 기존의 XML 질의어인 XQuery에 공간 SQL을 확장하여 GQuery를 제안하였다. 본 논문에서 제안한 GQuery는 사용자에게 GML 문서에 대한 통합된 질의 환경과 유연성을 제공해 준다.

향후 과제로는 본 논문에서 제안한 GQuery를 GML 문서에 대한 통합 검색을 지원하는 미디어이터 시스템 구현에 적용하는 것이다.

참고 문헌

- [1] OGC, Geography Markup Language (GML) Implementation Specification 3.0, 2003
- [2] OGC, Conformance Test Guidelines for OpenGIS Simple Features Specification for SQL, Revision 1.0. 1998.
- [3] W3C, XML Specification ("XMLspec"), 1998
- [4] W3C, XQuery 1.0: An XML Query Language, 2002.
- [5] W3C, Xml Path language(XPath) 1.0, 1999.
- [6] Angela Bonifati, Stefano Ceri, "Comparative Analysis of Five XML Query Languages" SIGMOD, Vol.29, NO. 1, March 2000.
- [7] Don Chamberlin, Jonathan Robie, Daniela Florescu, "Quilt: an XML Query language for heterogeneous data sources" LNCS, 1997.
- [8] 이환재, 안준순, 강동재, 이경모, 정보홍, 박동선, 배해영, "GEO/Millennium: 클라이언트-서버 공간 데이터베이스 시스템", 한국정보과학회 학술발표논문집, 27권1호, p.48- 50, 2000.