

# 5J-02 拡張可能 XML 問合せ言語 X<sup>2</sup>QL 処理系の開発

川田 純<sup>†</sup> 品川 徳秀<sup>††</sup> 北川 博之<sup>†††</sup> 石川 佳治<sup>†††</sup>

<sup>†</sup> 筑波大学 システム情報工学研究科 <sup>††</sup> 筑波大学 工学研究科 <sup>†††</sup> 筑波大学 電子・情報工学系

## 1 はじめに

近年、各種文書データベースの構築やその利用の高度化が急速に進みつつある。特に、インターネットの著しい普及に伴い、XML をはじめとする構造化文書データベースの構築とその利用の高度化の要求が増大している [1]。XML データベースに対する基本的な処理記述言語として、これまでに各種の問合せ言語が提案されている [2][3]。

XML データベースに対する処理記述では、タグの階層構造で与えられる文書構造に対する処理と文字列テキストで与えられる記述内容に対する処理の、両者を扱う必要がある。前述の問合せ言語の多くは、高水準な宣言的記述による高い問合せ記述能力を持っており、特に XML-QL[2] や半構造データ問合せ言語のいくつかは文書構造に対する柔軟な再構成機能を持つ。

一方で、これらは記述内容処理に関しては極めて限定された機能しか提供していない。また、構造操作に関しても、宣言的に記述する事が困難な応用処理も存在する。このような場合には、より詳細かつ具体的な記述能力が要求される。

我々はこのような問題に対し、XML 文書に対する問合せ言語が持つべき文書構造操作機能に加え、ユーザ定義の外部関数による機能拡張性を備えた拡張可能 XML 問合せ言語 X<sup>2</sup>QL を提案し、開発を進めている [4]。

本稿では、大量の XML 文書を格納管理する XML サーバ機能と連携した X<sup>2</sup>QL 処理系の開発について報告する。本 X<sup>2</sup>QL 処理系では XQL 問合せを処理可能な商用 XML サーバをそのコンポーネントとして利用する。従って、X<sup>2</sup>QL 問合せが与えられた際 XML サーバから必要な XML 文書を抽出するため、XQL に基づく絞込み問合せを発行する。本稿ではこの手法についても言及する。

## 2 拡張可能 XML 問合せ言語 X<sup>2</sup>QL

X<sup>2</sup>QL 問合せは、XML-QL をベースとした構文によって記述される。また、問合せ中には外部プログラムとして与えられる外部関数が利用可能であり、これによって高い機能拡張性が要求される処理に対応する事が可能である。2.1 節で X<sup>2</sup>QL 問合せについて説明し、2.2 節で外部関数について述べる。

### 2.1 X<sup>2</sup>QL 問合せ

X<sup>2</sup>QL 問合せの基本構文を次に示す。

```
where パターン式および変数束縛 [in 対象]
      [, パターン式および変数束縛 [in 対象]]* [, 述語]*
[rank-by 順位付け基準式 top 選択件数]
[order-by 出力順序基準式 [descending]]
construct 出力生成式
```

このうち、where は問合せ条件として適合すべき文書構造やそれが満たすべき条件を指定し、必要に応じて変数の束縛を指定する節である。変数はその記述位置

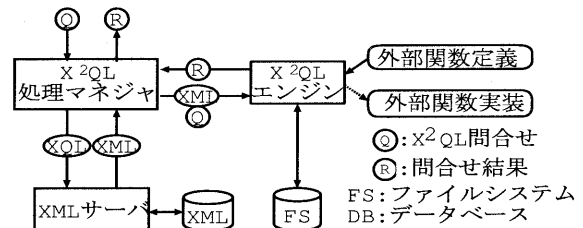


図 1: システム概略図

に応じて、文書要素、文書要素内容、文書要素名、属性名、属性値のいずれかに束縛され、対応する述語によってその値の満たすべき条件が記述される。construct 節は where 節のパターン式中で束縛された変数のそれぞれの組合せに対して生成される出力の構成を指定する節である。rank-by、order-by 節について本稿では説明を省略する。X<sup>2</sup>QL の構文は基本的に XML-QL の構文に従うが、後述のとおり外部関数が適宜利用可能である。

### 2.2 外部関数

高度な文書処理においては、必要とされる機能が多岐に渡るため、組込み関数を網羅的に提供する事は本質的に不可能である。そこで、X<sup>2</sup>QL においては、Java 等の汎用プログラミング言語による外部プログラムとして外部関数を与える事を可能とし、問題の解決を図っている。

外部関数は、不特定の対象に適用可能な一般関数と、特定の文書要素を持つ文書要素名に対して限定的に適用される文書要素固有メソッド (以下、メソッド) とに分けられる。文書要素毎に異なる処理を同名のメソッドとして定義する事も可能である。ユーザは必要に応じて外部関数を導入する事で問合せに必要な機能を拡張する事ができる。

## 3 X<sup>2</sup>QL 処理系の概要

現在、論文 [4] に述べたアプローチに基づいて X<sup>2</sup>QL 処理系の構築を行っている。本システムの概略図を図 1 に示す。システムは、XML サーバ、X<sup>2</sup>QL 問合せ処理エンジン (X<sup>2</sup>QL エンジン)、X<sup>2</sup>QL 処理マネージャから構成される。XML サーバは大量の XML 文書を格納管理し、XQL [3] に基づく XML 文書データの問合せ機能を提供する。本 X<sup>2</sup>QL 処理系では、この XML サーバの機能を用いて、与えられた X<sup>2</sup>QL 問合せが対象とする XML 文書の効率的な絞込みを実現する。X<sup>2</sup>QL エンジンは、XML サーバにより抽出された XML 文書データに対し、X<sup>2</sup>QL 問合せの処理を適用する。X<sup>2</sup>QL エンジンは必要に応じて外部関数の呼出しも行う。X<sup>2</sup>QL 処理マネージャは前二者を統合し、問合せ処理の流れを制御する。与えられた X<sup>2</sup>QL 問合せに基づいて XQL 問合せを生成し XML サーバから処理対象の XML 文書を抽出し、X<sup>2</sup>QL エンジンの呼出しを行う。

XML サーバを 3.1 節で、また X<sup>2</sup>QL エンジンを 3.2 節、X<sup>2</sup>QL 処理マネージャを 3.3 節で述べる。

### 3.1 XML サーバ

本研究では、現在、XML サーバとして、ビーコン IT 社の XML サーバ Tamino [5] を使用している。Tamino の主な機能は、XML 文書の格納と、クライアントか

Design and Implementation of an X<sup>2</sup>QL Query Processing System  
Jun Kawada<sup>†</sup>, Norihide Shinagawa<sup>††</sup>, Hiroyuki Kitagawa<sup>†††</sup>,  
Yoshiharu Ishikawa<sup>†††</sup>  
<sup>†</sup> Graduate School in Sys. and Info. Eng., Univ. of Tsukuba  
<sup>††</sup> Doctoral Program in Eng., Univ. of Tsukuba  
<sup>†††</sup> Institute of Info. Sci. and Elec., Univ. of Tsukuba

らの格納 XML 文書の検索がある。Tamino は問合せ言語として、XQL のサブセットである Tamino XQL (4.1 節参照) を提供している。

前述のとおり、本 X<sup>2</sup>QL 処理系においては、XML サーバの検索機能を利用して XML 文書を効率的に検索する。

### 3.2 X<sup>2</sup>QL エンジン

X<sup>2</sup>QL エンジンでは、X<sup>2</sup>QL 問合せを XSLT<sup>†</sup> 記述に変換し、既存の XSLT 処理系を利用して問合せの処理を行う。X<sup>2</sup>QL エンジンについては文献 [4] に詳細を示している。

### 3.3 X<sup>2</sup>QL 処理マネージャ

X<sup>2</sup>QL 処理マネージャでは、XML サーバから問合せ処理の対象となる XML 文書を取得するために、X<sup>2</sup>QL 問合せをもとにして XQL 問合せを生成し、これを用いて XML サーバに対する問合せを行う。

次に、この XML サーバから返された検索結果である XML 文書と与えられた X<sup>2</sup>QL 問合せを X<sup>2</sup>QL エンジンに渡し、X<sup>2</sup>QL エンジンで処理を行う。

最後に、X<sup>2</sup>QL 処理マネージャは、この X<sup>2</sup>QL エンジンでの処理結果を問合せ結果として返す。

## 4 XQL を用いた絞込み問合せの生成

本節では、X<sup>2</sup>QL 問合せから XML サーバに対する XQL による絞込み問合せを生成する手順について述べる。

### 4.1 XQL

XQL では要素の条件や要素間の関係条件をパス表現で表し、問合せはその条件に合致する文書要素などの抽出を行う。尚、問合せ演算子として、フィルタ “[ ]”、親子関係 “/、//”、属性を表す “@” 比較演算子 “=、<、>、<=、...”、論理演算 “\$and\$”、“\$or\$” などを利用する事ができる。

XQL の記述の例を次にいくつか示す。

- (1) bookstore 文書要素の子孫要素である book の検索例  
bookstore/book
- (2) bookstore 文書要素のいずれかの子孫要素である title 文書要素の検索例  
bookstore//title
- (3) last-name 文書要素に Bob という内容をもつ全ての author 文書要素の検索例  
author[lastname = 'Bob']
- (4) from 文書属性に Harvard という内容をもつ全ての author 文書要素の検索例  
author[@from = 'Harvard']

なお、今回 XML サーバとして利用している Tamino が提供する Tamino XQL は、1998 年 9 月に W3C へ提出された XQL のサブセットにいくつかの付加機能を追加したものである。従って以下で述べるように、Tamino XQL の持つ制限を考慮して絞込み問合せを生成する必要がある。

### 4.2 絞込み問合せ生成の概要とその適用手順

XQL による絞込み問合せ生成を考える上で、与えられる X<sup>2</sup>QL 問合せ中には複数文書の結合条件の指定は無いものと仮定する。また、本研究では、等値比較

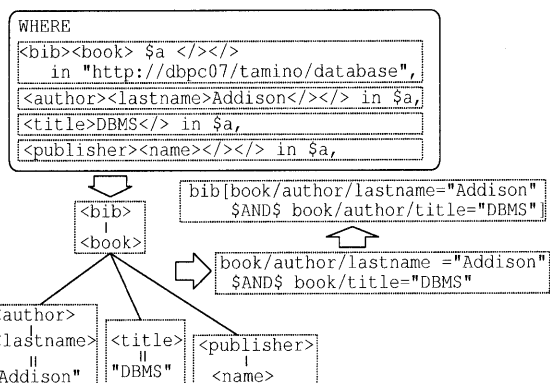


図 2: 絞込み問合せの生成

以外の述語と接続以外の正規表現は、絞込み条件としては考慮しない事とする。

X<sup>2</sup>QL 問合せの where 節に基づき Tamino XQL 問合せを生成する手順を次に述べ、例を図 2 に示す。

1. X<sup>2</sup>QL 問合せ中の where 節の条件を “;” で区切られた個々の条件 (condition) に分割する。
2. condition 内で in の右側に URI が指定されている condition をその依存関係を表す木構造のルートノードとする。
3. condition 内で変数が in の右側に現れるものが in の左側に現れるものの子ノードとする。
4. condition の末端ノードがリテラルと等値比較されているものを見つけてルートノードの子ノードからの文書要素のパスをそれぞれ \$AND\$ で結合する。
5. ルートノードに先の \$AND\$ で結合した条件を [ ] によるフィルタとして付加する。

## 5 まとめ

本稿では、XML サーバ、X<sup>2</sup>QL 処理マネージャ、X<sup>2</sup>QL エンジンから構成される X<sup>2</sup>QL 処理系の概要について報告した。

現在、本 X<sup>2</sup>QL 処理系を実装中である。今後、XML サーバへの絞込み問合せ生成に伴う X<sup>2</sup>QL 問合せの書き替えについても検討する予定である。

## 参考文献

- [1] World Wide Web Consortium, <http://www.w3.org/>.
- [2] A. Deutsch, M. Fernandez, D. Florescu, A. Levy and D. Suciu. A Query Language for XML, *Proceedings of the Eighth International World Wide Web Conference (WWW8)*, Computer Networks, Vol. 31, No. 11-16, pp. 1155-1169, 1999.
- [3] J. Robie, J. Lapp and D. Schach. XML Query Language (XQL). *The Query Languages Workshop (QL'98)*, <http://www12.w3.org/TandS/QL/QL98/pp/xql.html>, 1998.
- [4] Shinagawa, Kitagawa, and Ishikawa, X2QL: An eXtensible XML Query Language Supporting User-defined Foreign Functions, *Proc. 2000 ADBIS-DASFAA Symposium on Advances in Databases and Information Systems*, Prague, Czech, September 2000. (to appear)
- [5] Tamino ホームページ, <http://www.beacon-it.co.jp/products/inet/tamino/index.html>

<sup>†</sup>XSLT は W3C で策定中の XSL (XML スタイルシート言語) で用いられるルールベースの XML 変換言語である。