

6W-5 Quiz: 分散環境を想定した XML 問い合わせ言語

横山 昌平[†] 一色 誠司[†] 太田 学[‡] 石川 博[‡]

[†] 東京都立大学工学部電子情報工学科

[‡] 東京都立大学大学院工学研究科

1. はじめに

XML は EC 環境の拡大と共に、BtoB 間でのデータの交換手段として広く普及している。一方で、XML-QL や XML-Query^[1]等の様に、XML 文書をデータベースと見立て、それに問い合わせをする研究^[2]もまた盛んである。XML データはインターネット上、様々なサイトに分散して存在するが、それら XML データ同士の関係はほとんど無い。つまり、それぞれのサイト管理者の様式に従った XML 文書が、あるサイトではファイルとして、またあるサイトでは RDB に格納されて存在している。本稿では、これらインターネット上に分散している XML データの結合・問い合わせ、あるいは交換、処理の手続きモデル(ビジネスモデル)の記述言語である Quiz、実行システムの sQuiz を提案する。

2. Quiz で使用する諸技術

2.1 xGarden

プログラム上で XML 木構造を扱う方式には DOM がある。DOM のみでは XML 木構造の集合を扱う事が出来ない。従って独自形式を定義した。

xGarden は xField とメタデータからなっている (Fig.1 左)。xField は XML 木構造を複数格納することが可能である。xField は以下の4つの状態をとる。

- xForest: 複数の木構造
- xTree: ひとつの木構造
- xFoliage: 複数の値
- xLeaf: ひとつの値

2.2 xScissors

XML データ中の特定ノードや値を指し示す仕組みとして xScissors を提案する。これは Xpath を基に、Xpath 自身が持つ問い合わせ機能等を省略し、後述する '%' 修飾子の機能を付け加えたものである。以下の単純な XML データを例に説明する。

```
<LIST>
```

```
<BOOK year="1595">Romeo and Juliet </BOOK>
```

```
<BOOK year="1596">King Richard II </BOOK>
```

```
</LIST>
```

- LIST の子 BOOK 要素を指定する
/LIST/BOOK

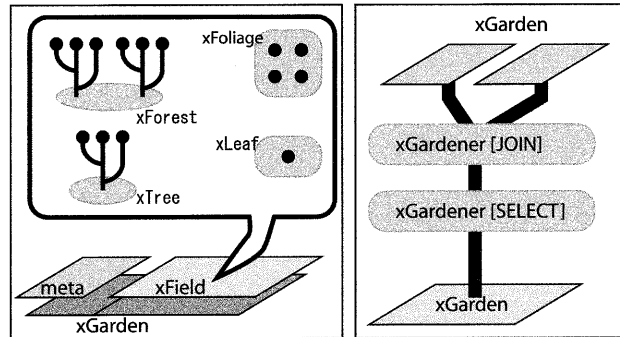


Fig.1 xGarden と xGardener

- BOOK 要素を指定する (階層の位置が不問の場合)
//BOOK
- BOOK 要素に含まれる文字列を指定する
/LIST/BOOK%
- year 属性を指定する
/LIST/BOOK/@year
- year 属性に属する文字列を指定する
/LIST/BOOK/@year%

3. Quiz

Quiz の文法は XML に従う。対象となるデータは xGarden であり、ノードの指定には xScissors を用いる。本稿では Quiz の機能として①SELECT、②JOIN そしてユーザーが手続きを定義する③外部関数を紹介する。全ての問い合わせは入出力に xGarden 形式をとる。これは Quiz の問い合わせ文(xGardener と呼ぶ)を多段に重ねられることを表している (Fig.1 右)。この性質により、我々は、一つの問い合わせが持つ機能を最小限にとどめ言語仕様の簡略化をはかり、尚かつ複雑な問い合わせには、xGardener を複数ネストすることで目的を果たせるようにした。そしてこの xGardener の組み合わせ方の違いはすなわちビジネスモデルの違いと考えることができる。

3.1 SELECT 文

```
<select return="/LIST" domain="/LIST/BOOK">  
  <from><xGarden src="http://~"></from>  
  <where>  
    <eq>  
      <argument v="1595">  
      <argument x="/LIST/BOOK/@year">  
    </eq>  
  </where>  
</select>
```

Quiz: An XML query language in a distributed environment
Shohei Yokoyama[†] Seiji Isshiki[†] Manabu Ohta[‡] Hiroshi Ishikawa[‡]
[†] Faculty of Engineering, Tokyo Metropolitan University
[‡] Graduate School of Engineering, Tokyo Metropolitan University

この Quiz 文は前述した XML のサンプルデータから BOOK 要素の year 属性が 1596 のものを返す問い合わせの例である。

木構造に対する問い合わせと RDB に対する問い合わせの違いは、WHERE 句が真の時、その影響範囲が一意では無いということである。つまり RDB であるなら、WHERE 句が偽となるタプルを全て無視できるが、階層構造を持つ XML の場合、どのノード以下を偽と見るか指定する必要がある。そこで domain という概念を用いる。domain で指定されたノードを繰り返す要素と考え、WHERE 句が偽の場合、domain まで遡り偽となったノードを無視する(Fig.2 左)。例では、Romeo and Juliet が偽であるので、domain で指定されている /LIST/BOOK 要素以下を全て無視する。また select 要素の return 属性は結果を返す最上位ノードを指す。よってこの問い合わせの結果は以下ようになる。

```
<LIST>
  <BOOK year="1596">King Richard II</BOOK>
</LIST>
```

3.2 JOIN 文

異なる XML ソース同士の結合は join 文で行う。以下に簡単な例を記す。

```
<join return="/LIST">
  <from domain="/ele1/ele2" root="/ele1/ele2">
    <xGarden src="http://~">
  </from>
  <to domain="/LIST/BOOK" mount="/LIST/BOOK">
    <xGarden src="http://~">
  </to>
  <requirement type="equality"
    from="/ele1/@year%"
    to="/LIST/BOOK/@year%" />
</join>
```

これは from 要素で指定される XML データ (join 元) を to 要素で指定される XML データ (join 先) に結合する問い合わせの例である。各 domain 属性に指定された繰り返し要素を総当たりで検索し、requirement 要素で指定されている条件に合致した組み合わせを結合する。join 元は root 属性を根に、join 先の mount 属性で指定されているポイントに接ぎ木される。(Fig.2 右)

3.3 外部関数

xGarden を別プロセスに渡し、結果を xGarden 形式で受け取る事を外部関数と呼ぶ。別プロセスに渡すプロトコルとして HTTP を用いている為、別ドメインのサーバーに渡す事も可能である。

```
<outer-function href="http://~">
  <from><xGarden src="http://~"></from>
  <argument><item name="n" value="v"></argument>
</outer-function>
```

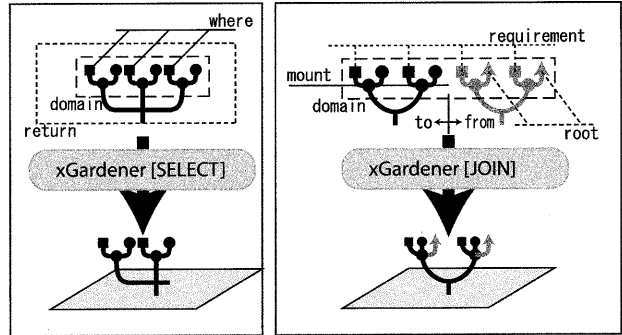


Fig.2 SELECT と JOIN

href 属性で外部関数の在処を指定し、それに渡す xGarden を from 要素で指定、外部関数の引数を argument 要素で指定している。

4. sQuiz

Quiz を処理する実行系を sQuiz と呼ぶ。これは PHP 言語により WEB サーバー上のサーバーサイドスクリプトとして動作している。また外部関数実行の際、HTTP を用いて xGarden のやりとり等の制御を行っている。また問い合わせ文内のユーザー変数を見つけだし、動的に HTML フォームを作成し、ユーザーに入力を促すことも行う。問い合わせの最適化として、他サイトの XML データに問い合わせを発行する際、sQuiz は対象サイトに sQuiz システムが存在するかを探る。もし存在した場合、xGardener を対象サイトに投げ、処理をそのサイトに委託し結果を xGarden として得る。これは、一般に元データの量に比べ、問い合わせ結果のデータ量が小さいことから、トラフィック減少を目的とした仕組みである。

5. まとめと今後の課題

本稿にてインターネット上に分散して存在する XML データに対し、選択・結合等を組み合わせてビジネスモデルを記述する言語 Quiz の提案を行った。そして外部関数という仕組みにより、単純な問い合わせ以外の処理を可能とするフレームワークを示した。また Quiz の実行系である sQuiz についても紹介した。

今後の課題として、sQuiz での XML データのやりとりに SOAP³⁾を採用したいと考えている。また、Quiz 自体の機能拡張として、XML で表現された画像データ同士の結合を JOIN 文で行える様にしたいと考えている。また他の XML 問い合わせ言語を Quiz に変換するなど、Quiz の持つ拡張可能性をさらに発展させ、Quiz をビジネスモデル記述方法の基盤としたい。

参考文献

- [1] The XML Query: <http://www.w3.org/TR/query-algebra/>
- [2] H.Ishikawa, et al.: Querying Web Distributed Databases for E-Businesses: Requirement Analysis, Design, and Implementation, ADC2001 (2001)
- [3] SOAP: <http://www.w3.org/TR/SOAP/>