

例示に基づく XML 操作系における問合せ文生成機構

1Z-05

毛利隆軌[†] 松本明^{††} 森嶋厚行^{†††} 北川博之^{††††}[†]筑波大学 第三学群 情報学類 ^{††}筑波大学 システム情報工学研究科^{†††} 芝浦工業大学 工学部情報工学科 ^{††††} 筑波大学 電子・情報工学系

1 はじめに

近年、様々な情報源に格納されているデータを Web ページとして出力することが重要なデータ処理の一つとなっている。また、情報源として XML の重要性が増大している。我々は、XML 上の Web ビュー構築を支援するためのツール AQUA の開発を行っている [1]。AQUA の特徴の一つは、ユーザが XML データの一部を用いて Web ページの例を示すと、システムが情報源全体に対する Web ビュー構築を行う問合せを推論し実行する事である。その結果、Web ビュー構築に必要なデータ抽出操作とページレイアウト操作の融合を実現している。それに対し、他の Web サイト構築ツールではこれらの操作は全く別の操作である。例えば、データ抽出は SQL で、レイアウト操作は GUI でという風に行う。我々の実験では、AQUA を用いれば SQL 等を全く知らないユーザでも XML 上に Web ビューを構築出来ることが示されている。

現在までの AQUA システムでは、Web ビュー構築のための XML 操作を実行するために専用のエンジンを利用していた。一方、XML の問合せ言語としては XQuery が標準となりつつあり、様々な XQuery 処理系が現れてきている。さらに、問合せ最適化などの技法が発展するにつれ、ますます効率の良い XQuery 処理系が現れる事が予想される。そのため、AQUA での XML 操作を汎用の XQuery 処理系で処理できれば、より汎用性・実用性の高い Web 構築ツールとなると考えられる。本稿では、AQUA の概要、および AQUA における Web ビュー構築のための XML 操作を XQuery として出力するための仕組みについて説明する。

2 AQUA の概要

図 1 に AQUA のアーキテクチャを示す。AQUA の画面は、情報源である XML 文書集合を表示させるための DataBox(a) と、生成する Web ビューのテンプレートを表示するための Canvas(b) の 2 つの Window から構成される。DataBox に表示されている XML 文書から、操作対象となる要素や値 (以下オブジェクト) を Canvas に Drag&Drop(以下 D&D) することで、Web ビューを構築することができる。ユーザが DataBox 中のオブジェクトを Example と指定すると、そのオブジェクトに対する操作はそのオブジェクトが代表とするオブジェクト集合に対する操作と解

釈される。この集合を TargetSet と呼ぶ。またあるオブジェクトを Another と指定するとそのオブジェクトも含むように TargetSet が広がる。Clue と指定するとそのオブジェクトに関する選択条件で TargetSet を小さく出来る。複数の Example オブジェクト間には関連がある。相対的な位置関係により決まる関連を S-Association と呼び、オブジェクトの値が同一である場合に決まる関連を V-Association と呼ぶ。

DataBoxManager(c) は DataBox を管理し、TargetSet と Association の情報を導出する。CanvasManager(d) は Canvas を管理し、grouping 情報および layout 情報を導出する。QueryConstructor(e) で XQuery 問合せを生成し、XQueryEngine に実行させる。SiteConstructor(f) は、結果として得られる XML 文書から、HTML や SMIL ページ群を生成する。

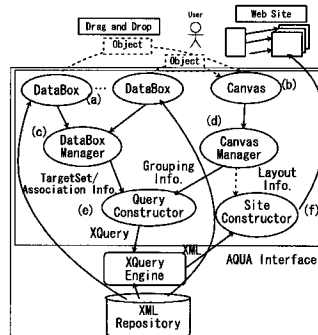


図 1: AQUA の概要

3 操作例

本節では具体的な例 (図 2) を用いて AQUA の操作を説明する。DataBox1(D1) には研究者の個人情報 (名前、年齢など) の XML 文書が表示されており、DataBox2(D2) には書誌情報をもつ XML 文書が表示されているとする。このとき「年齢が 25 歳以上の著者についてその著者毎の論文タイトルリストの Web ページを作成する。インデックスページには著者名を表示し、論文リストへのリンクを示す。」という操作を行いたいとする。その時の手順を以下に示す。

1. Canvas に HTML 用スペースを 2 つ表示させる。
2. D2 の “Suzuki” を Example と指定する。次に “Sato” を Another 指定する。システムは、Suzuki, Sato を含むすべての author がこの Example の TargetSet に入ると推論する。
3. D1 の “Suzuki” を Example と指定する。2. で指定した Example と値が同一なので、両者の TargetSet 間には V-Association が生じる。V-Association とは TargetSet を等結合することを意味する。
4. D1 の “26” を Clue と指定する。Clue の条件を

Query Generation from Example-based XML manipulations for Web-views
Takanori Mouri[†], Akira Matsumoto^{††}, Atsuyuki Morishima^{†††}, Hiroyuki Kitagawa^{††††}

[†]College of Information Sciences, Univ. of Tsukuba

^{††}Graduate School of Sys. and Info. Eng., Univ. of Tsukuba

^{†††}Dept. of Info. Sci. and Eng., Shibaura Inst. of Tech.

^{††††}Institute of Info. Sci. and Elec., Univ. of Tsukuba

[≥25]とする。この操作によって年齢25歳以上の authorのみが TargetSet に含まれる。

5. D1 の”Suzuki”を D&D する。
6. D2 の”XML”を Example として D&D する。
7. D2 の D&D した”XML”を repetition と指定 (* マークを追加)する。これにより同一著者の論文のタイトルが1つのページに集められる。
8. D1 の D&D した”Suzuki”を repetition と指定する。これにより著者名が1つのページに集められる。
9. HTML-1 上の”Suzuki”をアンカーとして HTML-2 へのリンクを作成する。

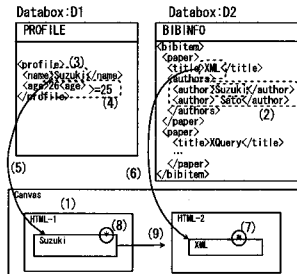


図 2: 論文リストを生成する操作例

4 Web ビュー構築のための XQuery 生成機構

AQUA のセマンティクスは、パス式と入れ子型リレーショナル代数によって定義されている [1]。本節では図 1 の QueryConstructor が Web ビュー構築のための XQuery 問合せを生成する仕組みを説明する。この問合せの結果には、Web ビューに必要な情報が現れるが、レイアウトに関する情報は含まれない(図 3)。レイアウト情報を用いて最終結果を作成するのは、SiteConstructor である。3 節の例では、図 6 の問合せが生成される。生成に使われるのは、DataBox-Manager から与えられた TargetSet と Association 情報(図 4)。3 節の操作 2 及び操作 4 で指定された TargetSet と、操作 3 での V-Association の例と、CanvasManager から与えられた Grouping 情報(図 5)である。それぞれの情報にある ID は、各 TargetSet を代表する Example の ID である。

```
<result>
  <item>
    <o1>Suzuki</o1>
    <rep>
      <o3>XML</o3>
      <o3>XQuery</o3>
    </rep>
  </item>
</result>
```

図 3: 生成される XML の一部

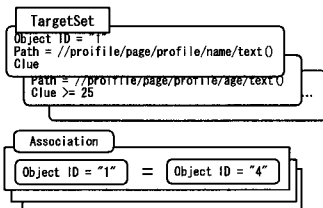


図 4: TargetSet/Association 情報

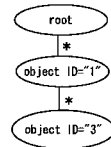


図 5: Grouping 情報

```
1. LET $TR:=<result>{
2. FOR $p1 IN //prof/page/profile
3. FOR $o1 IN $p1/name/text()
4. FOR $o2 IN $p1/age/text()[.>=25]
5. FOR $p2 IN //papers/page/paper
6. FOR $o3 IN $p2/title/text()
7. FOR $o4 IN $p2/author//text()
8. WHERE $o1=$o4
9. RETURN <page><n-cond><o1>{$o1}</o1><o2>{$o2}</o2>
   </n-cond><o3>{$o3}</o3></page>
10.}</result>
11.RETURN <result>
12.{
13. FOR $t IN distinct($TR/page//n-cond)
14. RETURN <item>{$t//o1}
15. <rep>{
16. FOR $ta IN $TR/page
17. WHERE $t=$ta//n-cond
18. RETURN $ta//o3
19. }</rep>
20. </item>
21.}
22.</result>
```

図 6: 3 節の操作例から生成される XQuery 問合せ

はじめに Association に基づいて各々の TargetSet を結合する。具体的には、TargetSet 情報と、Association 情報に含まれる S-Association 条件を基に FOR 文を生成する(2-7 行目)。次に V-Association を基に WHERE 節を生成する(8 行目)。さらに、RETURN 節を生成する。ここには、次の Nest 操作のために <n-cond> タグを埋め込む。

次に、grouping 情報の指示に従い Nest 演算に相当する処理を行う部分を生成する(13-20 行目)。<n-cond> タグは、Nest のための条件として利用される。データの重複を除去するため distinct 関数を利用する(13 行目)。以上の処理により XQuery 問合せが生成される。

5 まとめと今後の課題

本稿では、XML 上の Web ビュー構築支援ツール AQUA の概要、および、Web ビュー操作のための XML 問合せ生成機構について述べた。この機構により、汎用の XML エンジンのパフォーマンスを利用した Web ビュー構築が実現可能になる。今後の課題としては、静的 Web ビューのための問合せだけでなく、動的 Web ビュー生成をサポートするような XML 問合せの生成メカニズムの開発などが挙げられる。

謝辞

プロトタイプシステム構築にあたり多大なる助言を頂きました高野智氏に感謝いたします。本研究の一部は、日本学術振興会科学研究費補助金による。

参考文献

[1] A. Morishima, S. Koizumi, and H. Kitagawa. Drag and Drop: Amalgamation of Authoring, Querying, and Restructuring for Multi media View Construction Proc. 5th IFIP 2.6 Working Conference on Visual Database Systems (VDB5), 2000.