

7W-02 異種情報源統合環境におけるマルチメディア Webビュー構築ツールの設計と実装

高野 智[†] 小泉 清一^{††} 森嶋 厚行^{†††} 北川 博之^{†††}

[†]筑波大学 第三学群 情報学類 ^{††}筑波大学 工学研究科 ^{†††}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。それらのデータは、しばしば、画像、音声などを含むマルチメディアデータである。この処理においては、情報源から抽出されたデータを、どうレイアウトし表示するかといったページ設計が大きな問題となる。既存の枠組みのほとんどでは、情報源に対する問合せなどのデータ抽出操作とページ設計操作を異なる操作系で行っている。

本稿では、これらの操作をシームレスに統合した Web ビュー構築ツールの設計、実装について述べる。具体的には、情報源中のデータを Drag&Drop して配置することにより、情報源からのデータ抽出操作においても、通常の Web ページオーサリングツールと同様の操作感を提供する。本研究の特徴は次の 2 点である。(1) データオブジェクトを Example として指定することにより、通常のオーサリングツールと同様の object-at-a-time のページ設計操作と、set-at-a-time のデータ抽出操作を一体化している。(2) 情報源の異種性や半構造化に対応するため、Example が表すデータオブジェクトの集合を静的なスキーマ情報に依らずにグラフ構造をベースに動的に決定する仕組みを持つ。これによって、ユーザはマウスを用いたインスタンスベースの操作を行うだけで、システムが操作の意味を推論し、ユーザの要求にあった問合せを生成することを目的とする。

2 基本概念

本ツールの基本概念は次の通りである。

Object: D&D 対象の単位。構造化文書 (XML 等) のエレメント、及び RDB 中の値が該当する。

Window: 次の 2 種類がある。

- **DataBox:** 情報源中のデータ集合 (例えば Web ページやリレーションのタプルの集合) を表現する。画面には集合中の 1 つのデータ (例えば 1 つの Web ページや 1 タプル) のみが表示される。
- **Canvas:** ユーザはここに DataBox からオブジェクトをドロップし、欲しい情報を描画する。

Drag & Drop (D&D): DataBox から Canvas へオブジェクトを D&D し、配置することにより、様々なマルチメディア Web ビューの生成が可能となる。

Example: ユーザが、DataBox 中のオブジェクトを「Example である」と指定すると、そのオブジェクトの D&D 操作は、そのオブジェクトが代表となるオブジェクト集合に対する操作であると解釈される。

Target Set: Example オブジェクトが代表するオブジェクト集合。

Association: 異なる Target Set 中のオブジェクト間の関連を示す。Association には次の 2 種類がある。

- **S-Association:** 2 つの Example 間の相対的な位置関係により決まる。
- **V-Association:** Example として指定されたオブジェクトの値が同一である場合に決まる。

Design and Implementation of a Multimedia Web View Authoring Tool for a Heterogeneous Information Integration Environment

Satoshi Takano[†], Seiichi Koizumi^{††}, Atsuyuki Morishima^{†††}, Hiroyuki Kitagawa^{†††}

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

^{††}Doctoral Program in Eng., Univ. of Tsukuba

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

これらがあると、その関連をもつ特定の組み合わせだけが操作の対象となる。

3 基本操作

基本操作は、DataBox 中のオブジェクトを Canvas に D&D し、配置するだけである。ただし、DataBox においてマウスをオブジェクトの上で右クリックすると、[Example, Another, Clue] というメニューが出現するので、Drag 前にそのオブジェクトの役割を指定することにより、多様な操作が可能となる。操作手順を次に示す。

(Example 指定 (Another 指定 | Clue 指定)* | D&D)*

Example でないオブジェクトの D&D は、単にそのオブジェクトを Canvas に配置することを示す。Example と指定した場合は、TargetSet 中のオブジェクト群の操作を代表した操作と解釈される。最初はデフォルトの TargetSet が設定される。Another は、Another と指定されたオブジェクトも含まれるように TargetSet を広げる。オブジェクトを Clue と指定すると、そのオブジェクトの存在が、Example が TargetSet に含まれるための手がかり (条件) であることを示す。これは、TargetSet を狭める働きをする。Clue を指定するときには、そのオブジェクトが満たすべき条件をより詳細に記述することも可能である。Association は、暗黙に決定される。

4 操作例による解説

操作例を用いてより詳細な操作法を解説する。まず以下に述べる 2 種類の情報源があると仮定する。(1) Web 上に、選手リストを含む各チーム毎のページ (XML ページ) が存在し、そのページから各選手毎のページにリンクが張られている。(2) RDB (Video DB) には打席毎の動画オブジェクトとその打席の情報 (バッター名、ピッチャー名) を格納したリレーション VIDEO (VID, Batter, Pitcher, Contents) がある。このとき、「年齢が 20 歳以上の選手毎に打席の映像をまとめたビュー (SMIL ページ) を作成する。そしてインデックスページ (HTML ページ) を作成し、そこで各選手毎に選手名とビューへのリンクを示す」という例について考える。操作例の概要は図 1 に示す。

データ操作の元となる DataBox は図 1(a)~(c) のように示される。各 DataBox には既に操作対象となるデータ集合が集められていると仮定する。このデータ集合は DataBox の「next」「previous」ボタンでブラウジングできる。DataBox1 (TP) には各チーム毎のページ、DataBox2 (PP) には各選手毎のページ、DataBox3 (VD) にはリレーション VIDEO が格納されている。

操作手順を以下に示す。

1. リストを作成する HTML 用のテンプレートを用いて HTMLCanvas ウィンドウを開く。
2. TP の「Decker」を Example と指定する。このとき、DataBox 中の他ページにおいて「Decker」と同じ位置にあるオブジェクトの集合が TargetSet TS_1 に格納される。続いて「Spark」を Another と指定することで TS_1 は拡大され、結果的に「Decker」が代表する TargetSet は全選手名を含むことになる。¹ 操作中の TargetSet はハイライト表示され、ユーザはどのオブジェクトが TargetSet に含まれるかを確認することが出来る。
3. Example「Decker」を D&D する。

¹ TargetSet の具体的決定法は [1] で述べられている。

- HTMLCanvas 上のリスト項目を repetition と指定 (*マークを追加)する。これにより、TS₁ 中の選手名が1つのページに集められる。もしこの指定がなければ1つのページに1人の選手名を含んだページがTS₁中の選手数分だけ作成されることになる。
- VDの"Decker"をExampleとする。これら同一の値をExampleとするTargetSet間にはV-Associationが生じる。このAssociationはこれらのTargetSetを等結合することを意味する。
- PPの"Decker"をExampleとし、"23"をClueとする。Clueの条件を[≥20]とする、すなわち、年齢が20歳以上の選手名だけがTargetSetに含まれる。
- ビューを作成するSMIL用のテンプレートを用いてSMILCanvas ウィンドウを開く。
- VDの"Decker"をD&Dする。
- TPの①をExampleとし、D&Dする。
- VDの②をExampleとし、D&Dする。
- ③をrepetitionとする。これにより、各選手毎の映像が一つの連続映像にまとめられる。
- HTMLCanvas上の"Decker"をアンカーとしてSMILCanvasへのリンクを作成する。

最後に「Generate!」ボタンを押すと、Canvasで指定したHTML及びSMILのWebページが得られる。

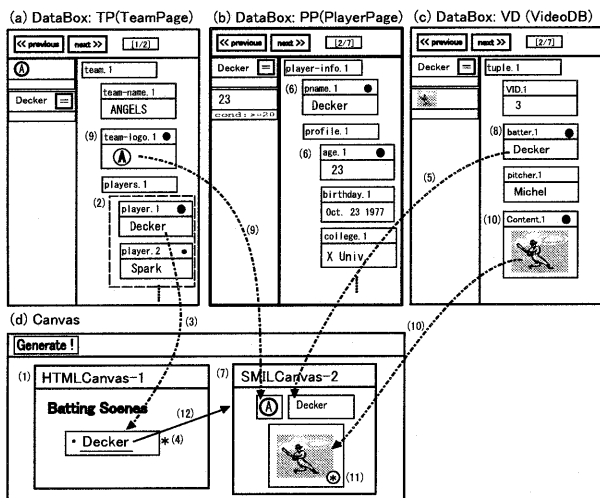


図 1: 操作例

5 プロトタイプシステムの実装

論文 [1] に述べられたアプローチに従って、本操作系のプロトタイプシステム「AQUA²」を実装した。なお、実装には Java2 SDK 1.3 を使用している。

図 2 にシステムの概要を示す。

システムは大きく 2 つの部分によって構成される。GUI 部はユーザーの操作を元に問合せを生成し、Engine 部がそれを解釈しメディアと通信を行い、データ抽出・再構成を行う。GUI 部と Engine 部間の通信は本プロトタイプシステムの本モジュールである AquaSystem がサポートする。

以下に各モジュールの機能を解説する。

GUI部:

GUI 部は、DataBox 部と Canvas 部から構成される。DataBox は DataBox を実現するモジュールであり、データ集合に含まれる 1 つのデータを表示する領域、指定された Example に関する情報をリスト表示する領域、データ集合をブラウジングするためのボタンを配置したツール

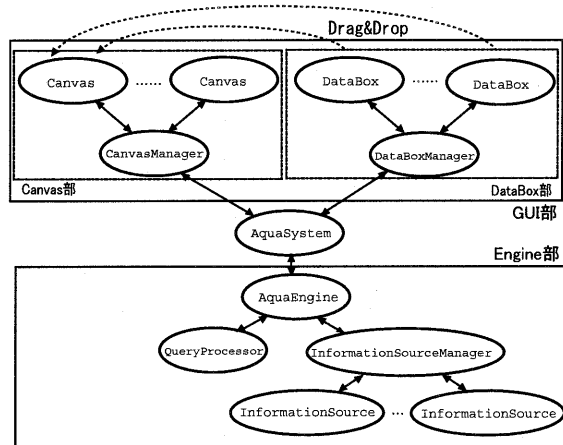


図 2: プロトタイプシステムの概要

バー、の 3 つのコンポーネントから成る。使用するデータ集合ごとに対応する一つの DataBox が生成され、ユーザーの指定に従って、TargetSet を導出するための部分問合せ式を生成する。DataBoxManager は DataBox を管理するモジュールで、複数の DataBox と通信し、部分問合せ式と Association から全体のデータを抽出するための問合せ式を生成する。

Canvas は Canvas を実現するモジュールであり、作成する Web ページの例ごとに一つ生成される。CanvasManager は Canvas を管理するモジュールであり、各 Canvas によって導出されるテンプレートと Canvas 間のリンク構造を元に、出力する Web ビューを構築するための再構築情報を生成する。

Engine部:

Engine 部は主に、問合せを処理するモジュールと (QueryProcessor) と、情報源へのアクセスを管理するモジュール (InformationSourceManager) に分けられる。AquaEngine はこれらを管理するモジュールであり、システムの下にある異種情報源統合環境を抽象化する。

AquaSystem:

AquaSystem は本システムの本モジュールであり、システム全体を管理する。操作例 (図 1) においてユーザーが「Generate!」ボタンを押すと、AquaSystem は DataBoxManager、CanvasManager のそれぞれと通信を行い、データ抽出のための問合せ式と再構築情報から、Web ビュー構築のための最終的な問合せ式を生成する。この問合せ式は AquaEngine を通じて QueryProcessor へと渡され、Web ビューが構築される。

6 まとめと今後の課題

本稿では、異種情報源統合環境におけるマルチメディア Web ビュー構築ツールの設計と実装について述べた。本ツールの実装に関しては、現在 Java Applet によるオンラインデモ³ が利用可能になっている。

本ツールの有効性については、XML-QL、XML-GL との比較評価実験により、使用感・所要時間における優位性が示されている。

今後の課題としては、1) よりユーザーの直観に沿ったデータ抽出機構の開発、2) より複雑な Web ビューが構築可能な Canvas の開発、などが挙げられる。

参考文献

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

² AQUA: Amalgamation of QUerying and Authoring

³ <http://www.dblab.is.tsukuba.ac.jp/kde/AQUA/>