

## 異種情報源統合利用環境 *Info Weaver*

### デモ 2

水口 弘紀<sup>†</sup> 梶野 智行<sup>††</sup> 小泉 清一<sup>†</sup> 森嶋 厚行<sup>†††</sup> 石川 佳治<sup>†††</sup> 北川 博之<sup>†††</sup>  
<sup>†</sup>筑波大学 工学研究科 <sup>††</sup>筑波大学 理工学研究科 <sup>†††</sup>筑波大学 電子・情報工学系

#### 1 はじめに

近年、インターネットをはじめとして、ネットワーク環境が急速に発展している。これにともない、ネットワークを介した様々な情報源へのアクセスが容易になりつつある。また、情報源として、伝統的なデータベースに加え、構造化文書や Web データが急増し、異種分散情報源の統合利用への要求は高まっている。

我々は、このような要求に対応するため、異種分散情報源統合利用環境 *Info Weaver* の研究開発を行っている。現在、*Info Weaver* では、Web、リレーショナルデータベース、構造化文書を具体的な統合利用対象としている。

*Info Weaver* は、統合データモデルとして我々が提案している WebNR/SD データモデルを採用し、メディアータ、ラッパーを用いたアーキテクチャで構成されている。

#### 2 *Info Weaver* の概要

図 1 に *Info Weaver* のアーキテクチャを示す。メディアータは各情報源にまたがった WebNR/SD データモデルに基づく統合的な操作を与える。それぞれの情報源に配置されたラッパーは、各情報源のスキーマ情報やコンテンツ情報などをメディアータに提供する。リレーショナルデータベースと構造化文書リポジトリに対するラッパーは、ローカルなそれぞれの情報源にアクセスする。また、Web に対するラッパーは、複数の Web サーバにアクセスすることで情報をメディアータに提供する。GUI はユーザからの問合せを受けとりメディアータにこれを渡し結果を表示する。本環境では、問合せは以下の手順にしたがって実行される。ユーザは、GUI を介して統合操作要求をメディアータに発行する。メディアータではこれを分析し、各ラッパーで処理可能な部分問合せに変換する。各ラッパーはそれぞれの情報源で処理可能な問合せに翻訳し情報源に投入する。ラッパーはその結果を WebNR/SD 統合データモデルに変換しメディアータにこれを返す。メディアータでは各ラッパーからの部分結果を基に最終的な統合要求を処理する。GUI ではこの結果をユーザに表示する。

*Info Weaver* : An Information Integration System on Heterogeneous Information Sources.

Hironori Mizuguchi<sup>†</sup>, Tomoyuki Kajino<sup>††</sup>, Seiichi Koizumi<sup>†</sup>,  
 Atsuyuki Morishima<sup>†††</sup>, Yoshiharu Ishikawa<sup>†††</sup>  
 and Hiroyuki Kitagawa<sup>†††</sup>

<sup>†</sup>Doctoral Degree Program in Eng., Univ. of Tsukuba

<sup>††</sup>Master's Degree Program in Sci. and Eng., Univ. of Tsukuba

<sup>†††</sup>Institute of Info. Sci. and Elec., Univ. of Tsukuba

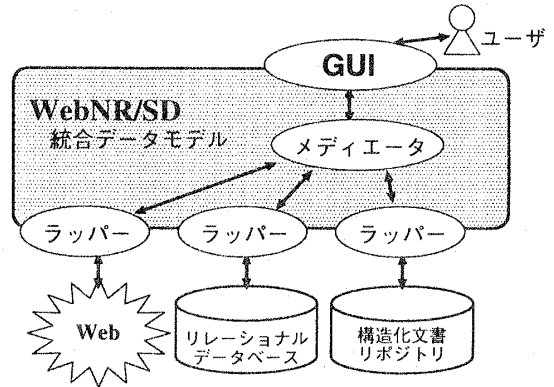


図 1. *Info Weaver*

#### 3 特徴

本アプローチには以下の特徴がある。(1) 対称的なデータ変換により、目的に応じて様々なデータ表現が可能である。例えば、本環境を用いることで、各種情報源を基にして、Web ビュー、リレーショナルビュー、構造化文書ビューを構築することができる。(2) 動的なデータ変換が可能である。これによって、異種情報源に対する統合操作における自由度が増す。例えば、構造化文書のデータを動的にリレーションに変換し、リレーショナル代数演算などを適用できる。(3) 異種情報の統合を目的とした視覚的ユーザインタフェース(図 3)。ユーザは、本視覚的操作系を用いて、各情報源の情報を閲覧することでデータ構造を把握し、これを基に視覚的に問合せを作成することが可能である。

#### 4 WebNR/SD

*Info Weaver* では、統合データモデルとして、我々が研究開発している WebNR/SD データモデル [1, 2] を用いている。WebNR/SD データモデルのデータ構造は、入れ子型リレーショナルモデルに構造化文書を扱うための抽象データ型 (SD 型) を導入したデータ構造である(図 2)。これにより、リレーション中では、構造化文書は SD 型の値として表現される。

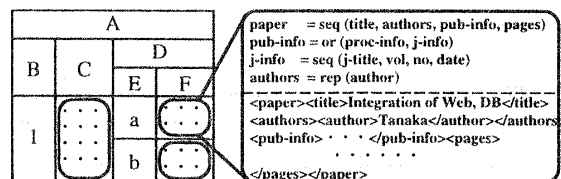


図 2. WebNR/SD のデータ構造

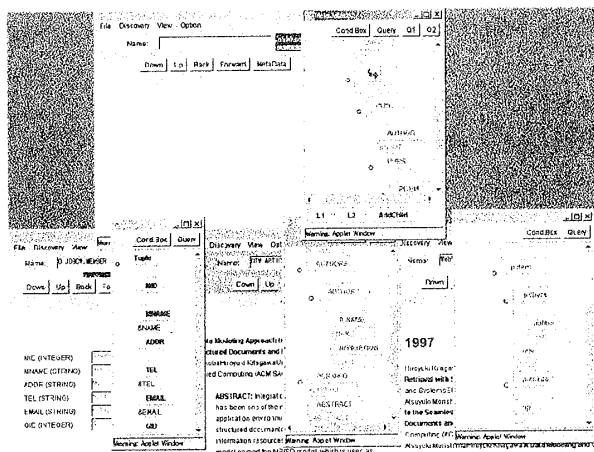


図 3. 視覚的操作系を用いた操作

WebNR/SD では、入れ子型リレーショナル代数と SD 値を扱うための固有の演算を提供する。これらに加えて、構造化文書と入れ子型リレーションを動的に相互変換するコンバータと、Web を操作する演算子群を持つ。これらの演算子群を用いて Web、リレーショナルデータベース、構造化文書に対する統合操作を行なう。

### 5 統合利用例

InfoWeaver を利用した統合操作例を示す (図 4)。

情報源として次の 3 つを考える。Web 上にデータベース研究室のホームページがある。ここには、データベース研究室で発表された年度別の論文リストページがある。それぞれのページには、発表した会議や掲載された学会誌へのリンクが含まれている (図 4(a))。構造化文書リポジトリには、発表された論文の要約と全文が格納されている (図 4(b))。リレーショナルデータベースには、データベース研究室の構成員情報が格納されており、メールアドレスや電話番号などが含まれている (図 4(c))。

この時、次のような要求があるとする。各構成員別の論文リストを Web ページとして作成したい (図 4(e))。それぞれのページには、タイトルだけでなく構造化文書リポジトリ中にある要約も掲載する。さらに、発表会議などへのリンクもあるものにしたい。また、リレーショナルデータベースにある構成員情報から構成員情報ページを作成し、作成した構成員別論文リストページへのリンクを含んでいるものにしたい (図 4(d))。

図 3 はこの統合操作を GUI を用いて行なっている例である。この統合は GUI を介してメディアータ、ラッパーの協調動作によって実現される。GUI における統合操作の記述は、視覚的操作記述言語 HQBE (Query By Example for Heterogeneous Information Sources) [3] を用いて行う。HQBE では、QBE と同様に、例示要素を用いて結果として必要なデータ構造を指定する。例示要素を指定することによって、データ同士の対応関係を記述する。

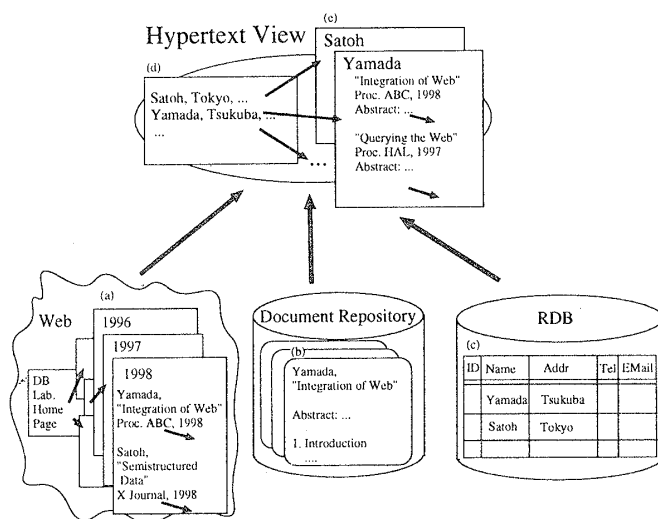


図 4. 統合操作例

### 6 おわりに

本稿では、異種情報源統合利用環境 InfoWeaver の概要を述べた。現在の実装は、構造化文書リポジトリ用の文献検索エンジンとして OpenText ならびに Namazu、リレーショナルデータベースとして Oracle8 ならびに mSQL を用いている。メディアータ、ラッパーは Java 言語を用いて開発を行った。また、モジュール間の通信には RMI を用いている。GUI は Java アプレットにより実装され Web ブラウザから実行可能である。

現在、MPEG などのマルチメディア情報も統合対象とするための研究開発を行っている。SMIL を統合の枠組とし、マルチメディアプレゼンテーションの構築を行うことが可能である。また、マルチメディアデータを統合するための新たな GUI も開発中である。さらに、PointCast などの Push 型情報を統合するための枠組も研究開発している。

### 参考文献

- [1] A. Morishima and H. Kitagawa. A Data Modeling and Query Processing Scheme for Integration of Structured Document Repositories and Relational Databases. Proc. 5th International Conference on Database Systems for Advanced Applications, pp. 145-154, 1997.
- [2] A. Morishima and H. Kitagawa. Integrated Querying and Restructuring of the World Wide Web and Databases. Proc. International Conference on Digital Media Information Base, pp. 262-271, 1997.
- [3] 森嶋厚行, 北川博之. 視覚的操作系による異種情報源統合利用環境支援. 電子情報通信学会論文誌 D-I, Vol. 39, No. 4, pp. 954-967, 1999.