

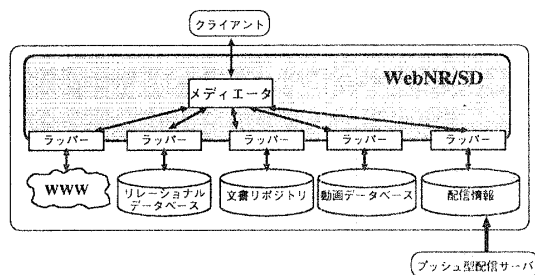
異種情報源統合利用環境 *Info Weaver* におけるプッシュ型情報配信4Q-8 水口 弘紀<sup>†</sup> 森嶋 厚行<sup>††</sup> 石川 佳治<sup>††</sup> 北川 博之<sup>††</sup><sup>†</sup>筑波大学 工学研究科 <sup>††</sup>筑波大学 電子・情報工学系

## 1 はじめに

ネットワークの発達により各情報源へのアクセスが容易になった。しかし、その中から欲しい情報を見つけ出すのが困難となりつつある。一方で、新しい情報は絶えず増え続けており、新鮮な情報が必要となっている。このような背景のもとに、クライアントに対してサーバが主体的に情報を送るプッシュ型の情報配信サービスが増加している。しかしながら、プッシュ型の情報配信サービスは、配信する側に制御が移るため、ユーザの希望しない情報も配信されることがある。また、複数の情報配信サービスの統合や、既存の情報源の情報とを組み合わせることによるユーザの好みに応じた情報配信等を実現することは重要であると考えられる。

一方、我々は、WWW、リレーショナルデータベース、構造化文書リポジトリ、マルチメディアオブジェクトを対象とした異種情報源統合利用環境 *Info Weaver* の研究開発を行なっている [1] [2]。 *Info Weaver* ではデータモデルとして WebNR/SD を用いている。我々が提案している統合データモデル WebNR/SD では、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と、WWW ページ間のリンク構造を扱うためのデータ型を導入している。さらに、構造化文書と入れ子型リレーション構造を相互変換する演算子と WWW ページ群を操作する演算子を持ち、構造化文書の再構成やハイパーテキストビューの構築が可能である。

本研究では、*Info Weaver* のもつ情報統合の枠組みをベースとしてプッシュ型情報源を含む各種の情報源を統合したプッシュ型情報配信について検討する。これによってユーザ要求に応じて情報を統合することが可能なメタプッシュ型情報配信サーバも実現することができる。

2 異種情報源統合利用環境 *Info Weaver*図 1. *Info Weaver*

*Info Weaver* は、情報源として WWW、リレーショナルデータベース、構造化文書、動画データベースなどのマルチメディア情報源を統合対象としていた。本研究では、プッシュ型の情報配信サービスを情報源としてさらに統合する (図 1)。各情報源にラッパーと呼ばれるモジュールを配置し、これにより各情報源の情報を統合データモ

Push type data dissemination in the information integration system *Info Weaver*.

Hironori Mizuguchi<sup>†</sup>, Atsuyuki Morishima<sup>††</sup>,Yoshiharu Ishikawa<sup>††</sup>, Hiroyuki Kitagawa<sup>††</sup><sup>†</sup>Doctoral Program in Eng., Univ. of Tsukuba<sup>††</sup>Institute of Info. Sci. and Elec., Univ. of Tsukuba

デル WebNR/SD に変換する。メディアータはこれに基づいた統合操作の機構を提供する。問合せがメディアータに投入されると、各ラッパーで処理可能な部分問合せに変換される。メディアータは各ラッパーの処理結果から最終結果を作成しユーザに結果を返す。今までの *Info Weaver* では、問合せが投入されるとすぐに演算が実行されていたが、本研究ではイベントの発生に対応して演算を実行する機構を導入する。

## 3 プッシュ型情報源の統合

インターネットを利用した PointCast、地上波を利用した bitcast、ADAMS などのプッシュ型情報配信サービスにおける特徴は、情報がユーザの側の操作によって配信されるのではなく、情報源から能動的に配信されることである。本研究では、このような動的な情報源に対応して、情報源に情報が到着した時、このイベントを受けて問合せが発行され、結果をユーザに能動的に配信する、イベント処理機構を導入する。また、本研究では、HTML のような構造化文書を配信するサービスを対象とする。

## プッシュ型情報源とラッパー

一つの配信サービスを一つのプッシュ型情報源とみなす。プッシュ型情報源におけるラッパーは、配信される情報のある単位を 1 タプルとしたリレーションビューをメディアータに提供する。配信サービスにおけるクライアントは情報の蓄積量に限界があり、サーバから配信された情報をすべて保持できるわけではない。そこで、このリレーションは格納可能なタプル数に制限があり、これをレシーバリレーションと呼ぶ。レシーバリレーションは情報の到着にしたがってタプル単位の追加削除が頻繁に行なわれる。

本研究では、リレーション  $R$  に対して、追加差分のタブルを表すデルタリレーション  $\Delta R$  を定義する。

$$r_2 := \sigma_{Title='DataBase'}(\Delta r_1) \quad (1)$$

(1) 式を一回行なうとレシーバリレーション  $r_1$  のうち、前回の操作以降に追加されたタブルが演算対象となり  $r_2$  には演算結果が入る。この時、演算対象となったタブルは消費され  $\Delta r_1$  リレーションにはタブルが存在しなくなる。その後、再度 (1) 式を行なうと、先ほど演算を行なった時点から再度行なうまでに、 $r_1$  に追加されたタブルの集合が  $\Delta r_1$  に対応し演算対象となる。本ラッパーはこのデルタリレーションもメディアータに対して提供する。さらに、情報が追加されたか、つまり、タブルの追加が起こったかを判別しイベントとして通知することができる。

## ルールに基づいたイベント処理

情報源への情報の追加を示すイベントをルールによって処理する。ルールは ECA ルールを用いて表現する。イベントとしてタブルの追加イベントを  $append(R)$  で表す。ここで、 $R$  は情報源におけるリレーションを指す。コンディションは  $|\Delta R| = N$  の条件を記述する。ここで、 $|\Delta R|$  はデルタリレーション  $\Delta R$  のタブル数を表し、 $R$  における追加の行なわれたタブル数が  $N$  であれば真となる。アクションとしては、WebNR/SD に基づく代数式を記述する。ルールは以下のように表現する。

$$\begin{aligned} \text{Rule} \{ & \text{Event} : \text{append}(r_3) \\ & \text{Condition} : |\Delta r_3| = 1 \\ & \text{Action} : \{r_4 := \sigma_{Title='DataBase'}(\Delta r_3)\} \} \quad (2) \end{aligned}$$

Rule{ Event : append( $r_4$ )  
 Condition :  $|\Delta r_4| = 1$   
 Action : {PUSH( $\Delta r_4$ )}} (3)

ここで、イベントの発生は図2のように起こるとする。空のリレーション  $r_3$  に3つのタプルが1回ずつ追加されると、イベント append( $r_3$ ) が3回発生する。ここでは(2)式がこのイベントに対応する。コンディションでは追加イベントに対応して  $|\Delta r_3|=1$  が真となり、アクションが実行される。ここでは、 $\Delta r_3$  は図中におけるIDが1のタプルのみを含む  $\Delta r_3$  リレーションを演算対象とする。この演算が終了すると、残りの2回の append( $r_3$ ) イベントが起き同様の処理が行なわれる。また、式(3)では、 $r_4$  へのタプルの追加が1個の時、これをクライアントへ配信することを示す。

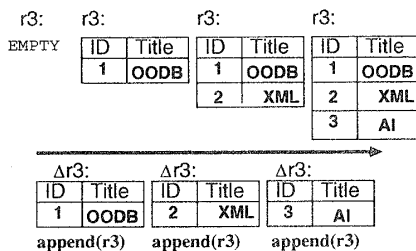


図2. イベントの発生タイミングとデルタリレーション

#### 4 InfoWeaverを用いたプッシュ型情報配信サーバの構築

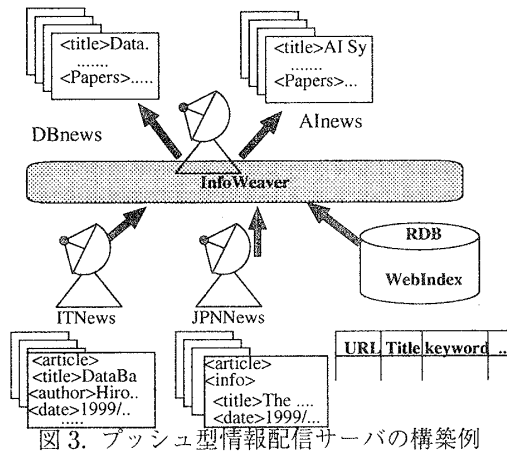


図3. プッシュ型情報配信サーバの構築例

情報技術ニュースを配信しているメールサービス IT-News と、Webを使ったニュース情報配信サービス JPN-News があるとす。また、リレーショナルデータベースには Web ページのインデックス情報が格納されており、URL とそれに対応するキーワードが格納されているものとする。

この時次のような要求が考えられる。ユーザは IT-News と JPNNews の2つのサービスから、データベースと人工知能に関する新しいニュース記事を作成し、それぞれをカテゴリとする2つの新しい配信サービスを作成したい。さらに、関連する情報の URL を記事とともに見せて欲しい(図3)。ITNews は新商品情報や各社の研究開発情報などが主でありユーザの注目度は高い。JPN-News は総合的なニュースであり、ニュース記事数が多いが注目度は低い。ここで、URL 情報データベースは図4のスキーマ、ITNews と JPNNews は図5、図6のDTD

を持っているとする。実際の操作は以下のように記述できる。まず、プロフィール情報を持ったリレーションを作成し、キーワードを指定しておく(図7)。つぎに3つのルールで新たな配信

サービスを記述する(図8)。はじめのルールはチャンネルの情報をつくり出すルールである。ITNews は注目度が高いため、追加記事が1つでもあればすぐにアクションを行なうよう、コンディションは  $|\Delta ITNews| = 1$  で記述する。アクションは演算が以下に行なわれている。(1)2つのチャンネル中の記事から必要な構造を抜き出す。(2)記事内容から探している記事のみを選択する。(3)Web インデックス情報データベースから Web ページのキーワードと記事内容の類似するものをさらに選択する。(4)リンク情報についての構造化文書を作成する。(5)新たな記事を作成、2つの新たなチャンネルを作成する。残り2つのルールでは作られた記事を配信するタイミングを決定し、記事が作られるとすぐに配信するように設定されている。

ID	URL	Title	Author	keywords
1		OODB		
2		XML		
3		AI		

図4. リレーション WebIndex

article	=	seq(title,authors,date,body)
authors	=	rep(author)
body	=	rep(p)

図5. ITNews DTD

article	=	seq(info,body)
info	=	seq(title,date,summary)
body	=	rep(p)

図6. JPNNews DTD

CID	ChannelName	Profile
1	DBnews	DB
2	AINews	AI

図7. リレーション ChannelInfo

```

Rule{ Event : append(ITNews)
    Condition :  $|\Delta ITNews| = 1$ 
    Action : {
        R1 :=  $\pi_{NewsTitle, Date, Body}(\Delta ITNews) \cup$ 
             $\cup_{Article \rightarrow (NewsTitle, Date, Body)}(\Delta ITNews) \cup$ 
             $\pi_{Title, Date, Body}(\cup_{Info \rightarrow (NewsTitle, Date)}($ 
             $\cup_{Article \rightarrow (Info, Body)}(\Delta JPNNews))$ 
        R2 :=  $\pi_{NewsTitle, Date, Body, ChannelName}(\cup_{Body \sim Profile} ChannelInfo)$ 
        R3 :=  $R_2 \bowtie_{Body \sim keywords} WebIndex$ 
        R4 :=  $P_{Ls(Link) \rightarrow Links: RC}(V_{Ls}=(Link)(WA_{(href:URL)} \rightarrow Link($ 
             $P_{(Title, Author) \rightarrow Link: SC, "a"}((R_3))))$ 
        R5 :=  $P_{(NewsTitle, Date, Body, Links) \rightarrow NewArticle: SC, "article"}(R_4)$ 
        DBnews :=  $\pi_{NewArticle}(\sigma_{ChannelName="DBnews"}(R_5))$ 
        AInews :=  $\pi_{NewArticle}(\sigma_{ChannelName="AINews"}(R_5))$ 
    }
```

```

Rule{ Event : append(DBnews)
    Condition :  $(|\Delta DBnews| = 1)$ 
    Action : {PUSH( $\Delta DBnews$ )}}
```

```

Rule{ Event : append(AInews)
    Condition :  $|\Delta AInews| = 1$ 
    Action : {PUSH( $\Delta AInews$ )}}
```

図8. 利用例における演算式

#### 5 まとめと今後の課題

本研究では、InfoWeaverへプッシュ型情報源を統合することで、複数の情報配信サーバを統合した情報配信や、他の情報源との統合的な情報のカスタマイズを実現した、メタプッシュ型情報配信サーバを構築した。その際、ルールによるイベント処理を導入することで、効果的な情報配信を行なうことが可能である。

#### 参考文献

- [1] A. Morishima and H. Kitagawa. InfoWeaver: Dynamic and Tailor-Made Integration of Structured Documents, Web, and Databases. *Proc. ACM Digital Libraries '99*, August 1999.
- [2] 水口弘紀, 森嶋厚行, 北川博之, "異種情報源統合利用環境におけるマルチメディアオブジェクト処理" 情報処理学会第57回全国大会, 1998年10月.