

配信型情報源統合環境における配信要求の視覚的記述方式*

2Z-01

岡本祥伯†

渡辺陽介††

石川佳治†††

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

††† 筑波大学電子・情報工学系

1 はじめに

近年、大量の情報源が利用可能となり情報統合への要求が増大している。またメールマガジンやデータ放送などの配信型情報源が多数利用されるようになってきている。我々の研究グループでは、配信型情報源を含む各種情報源を統合して扱うための機構の研究を行っている [1,2,3]。一般に、複数の配信型情報源同士や配信型情報源と他の情報源を統合して新たな情報配信サービスを提供することは容易ではない。我々は ECA ルールと呼ばれるイベントに応じた処理を行う枠組みを用いて、配信型情報源の統合を図っている。ユーザは ECA ルールを記述することによって、新たな情報配信サービスを定義することができる。しかし ECA ルールの記述は、ユーザにとって負担の大きいものであるため、よりユーザフレンドリーな配信要求の記述法が望まれる。そのための基盤として、我々はリレーショナル代数式による配信要求記述から ECA ルールを自動生成する枠組みを提案した [1,2,3]。しかし、一般のユーザに、さらに利用しやすいインタフェースが望まれる。そこで本稿では、リレーショナル代数式の枠組みに対する GUI として、QBE[4] をベースにした配信要求の視覚的記述法について述べる。

2 統合利用例

本節では配信型情報源を用いた具体的な統合例を示す。配信型情報源として、各企業に関するニュースをメールで配信している企業ニュース配信サービス (INews) と、各企業の株価の終り値の情報を Web 上で配信している株価情報配信サービス (SPrice) の 2 つがあるとする。企業ニュース配信サービスでは、企業名 (Name) とニュースのタイトル (Title) 及び記事内容 (Articles) が配信される。株式情報配信サービスでは、企業名 (CName) と業種 (Category)、その企業の株価の終り値 (Price) が 1 日 1 回配信される。本研究では配信型情報源をリレーションとして扱う。各配信サービスは、情報の到着時間を示す ITS 属性を持つものとする。またユーザの手元には、ユーザの所有株式リレーション (StockHolderInfo) があり、現在株を所有している各企業の企業名 (Company)、株の購入価格、株数、閾値 (Threshold) として注目している株価がリレーショナルデータベースに格納されている。

このとき要求として、「ユーザが株を所有する IT 関連企業の株価が指定された閾値以上であったら、その時の株価と企業名及び過去 2 日分のニュース記事を翌日の午前 0 時に配信して欲しい」という要求を考える。

3 統合アーキテクチャ

まず我々の研究グループがこれまでに開発してきた配信型情報源統合システムについて説明する (図 1)。ラッパーは、情報源のデータを受け取りそれをメディアータに渡す処理を行い、メディアータは情報の統合を行う。ルール処理モジュールは ECA ルールを保持しており、イベントの発生に応じて ECA ルールを評価する。ECA ルールとは、情報の

到着や時間の経過に起因して実行される処理を記述したルールである。その時行われる処理には、情報の蓄積、統合、配信などがある。また配信モジュールは、メディアータで生成された統合結果の配信を行っている。本研究で開発している配信要求の視覚的記述環境 EDGE(ECA rule Derivation using Graphical Environment) は、ユーザの視覚的操作から ECA ルールを生成するものである。GUI によってリレーショナル代数式を導出し、ルール生成モジュールによって ECA ルールを生成する。ルール生成モジュールの処理については [1,2] で述べられているので、本稿では GUI 部分を説明する。

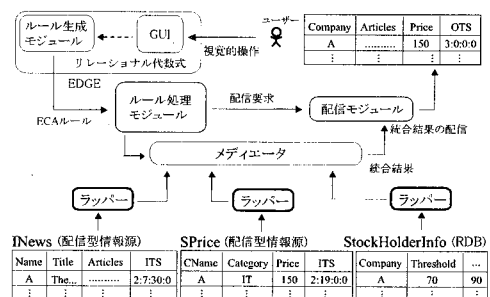


図1 統合アーキテクチャ

4 配信要求の視覚的記述

4.1 基本概念

本研究ではデータモデルとしてリレーションを用いている。リレーションに対する視覚的問合せ記述方式として、QBE (Query By Example) [4] が広く知られている。よってEDGEでは、QBEをベースにした時間情報に関する視覚的条件記述法を用いる。図2にGUIの構成要素を示す。それぞれの役割は以下の通りである。

RelationSkeleton 各リレーションの名前とITS属性以外の属性名が表示される。属性名の下に定数要素、例示要素、あるいはそれらを含んだ式を記述することで、その属性名に関する条件を指定する。

OutputSkeleton 統合結果のリレーションのスキーマを指定する。上段にユーザが任意に決定する属性名、下段に例示要素を記述する。

ConditionBox 論理演算子 $\wedge$, $\vee$ 等を用いた複雑な条件式を指定する。

TemporalConditionBox 配信スケジュール、配信内容の時間的な選択条件を指定する。時間軸上に各リレーションのITS属性の例示要素を配置して、時間的な前後関係を表す。表示されている横線は時間軸を表しており、縦線は同一の時刻を表す。配置されている記号について説明する。まず◎は統合結果の配信予定時刻を表しており、図中のOTSのことを指している。■はITS属性の例示要素を表している。原則として、1つの時間軸には1つのITS属性の例示要素しか配置できないものとする。□は中間要素を表しており、条件記述として時間の流れや幅を記述するときに使用する。図中で示されている矢印には、next, previous, after, beforeの4種類があり、中間要素間の関係を示すために使う。それぞれの役割については以下の通りである。

- 左から右への弧を描く矢印 (next) …時刻パターンにマッチするタイムスタンプのうち、時間的に開始点の次

*Visual Specification of Delivery Requirements for a Dissemination-based Information Intergration Environment

† Colledge of Information Sciences, University of Tsukuba

†† Graduate School of Systems and Information Engineering, University of Tsukuba

††† Institute of Information Sciences and Electronics, University of Tsukuba

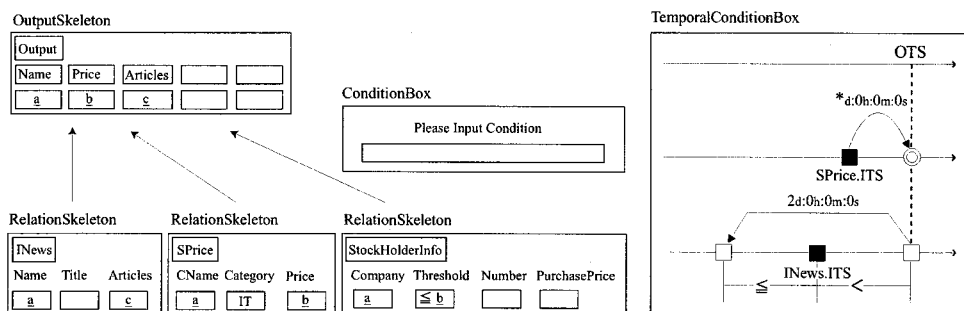


図2 操作例

の時刻を表すタイムスタンプを示す。

- 右から左への弧を描く矢印 (previous) …時刻パターンにマッチするタイムスタンプのうち、時間的に開始点の次の直前の時刻を表すタイムスタンプを示す。
- 左から右への角のある矢印 (after) …開始点の時刻から時刻パターン分だけ経過した時刻を表すタイムスタンプを返す。
- 右から左への角のある矢印 (before) …開始点の時刻から時刻パターン分だけ過去に遡った時刻を表すタイムスタンプを返す。

時刻パターンには次の形式のうちの1つを与えられる。

- (1) day:hour:minute:second
- (2) dayofweek:day:hour:minute:second

day, hour, minute, second の各フィールドに、非負整数かワイルドカード '*' を記述することができる。フィールド dayofweek には、文字列 {Sun, Mon, Tue, Wed, Thu, Fri, Sat} のうちの1つ、または '*' を指定できる。ただし previous と after には '*' を使うことはできない。

比較演算子 <, ≤ は、ITS 属性の例示要素と中間要素との比較関係を表したものである。

4.2 操作方法

このシステムの操作方法を、前述した利用例を用いて説明する。図2が利用例の条件を実際に記述したものである。配信要求の記述内容は以下の通りである。

1. ITS 属性以外の属性についての選択条件を RelationSkeleton に記述する。まず3つのリレーションの結合条件は、Name = CName = Company である。よって、それぞれの属性に共通の変数 *a* を記述する。次に IT 関連企業のニュース記事が必要なので、SPrice の Category 属性に "IT" を記述する。統合結果として、SPrice の株価と INews のニュース記事が要求されているので、それぞれの属性にあたる Price, Articles に変数 *b*, *c* をそれぞれ記述する。次に「株価が指定された閾値以上」という条件より、StockHolderInfo の Threshold 属性に Threshold ≤ Price を意味する「≤ *b*」を記述する。
2. 統合結果として必要なものは OutputSkeleton に記述する。今回必要となるのは、会社名とその会社のニュース記事と株価である。ここで RelationSkeleton で定義した変数 *a*, *b*, *c* を OutputSkeleton の下の段に記述する。上の段にはその変数に応じた属性をユーザが任意に記述する。
3. 次に TemporalConditionBox を立ち上げる。ここには時間的な条件を記述する。

利用例における時間的な条件を取り上げてみると、「SPrice のニュース記事が到着した次の日の午前0時に配信」かつ「過去2日分の企業ニュース」という部分が当てはまる。まず前者について考えてみる。初めに、1

つ目の時間軸上に ITS 属性を持つ情報源 SPrice を配置する。次に next を用いて、次の日の午前0時を示す矢印と配信予定時刻である OTS 属性を配置する。この時与える時刻パターンは *:0:0:0 となる。次に後者について、まず2つ目の時間軸に ITS 属性を持つ情報源 INews を配置する。次にそれを囲むように中間要素を2つ配置する。右側の中間要素は OTS 軸上に配置をして、その中間要素からもう片方の中間要素へ before を用いて過去2日を表す矢印を配置する。この時与える時刻パターンは 2:0:0:0 となる。また中間要素における時間幅の調整は比較演算子を用いて行う。今回の要求は「過去2日分」という時間幅要求なので、before の開始点には < を、終了点には ≤ を配置することで「過去2日」という時間幅を表現している。

5 まとめと今後の課題

本稿では、配信型情報源統合環境における配信要求の視覚的記述方法を述べた。このシステムにより、ユーザは容易に配信要求を記述することが可能となる。今後の課題としては、1) より複雑な問合せの記述、2) より多様な問合せの記述方法への対応などが挙げられる。

謝辞

本研究の一部は日本学術振興会科学研究費基盤研究 (B)(12480067) および文部科学省特定領域研究 (C)(13224008) による。

参考文献

- [1] 梶野 智行, 北川 博之, 石川 佳治, "配信型情報源に対する統合要求記述と ECA ルール生成", 電子情報通信学会論文誌.
- [2] Hiroyuki Kitagawa, Tomoyuki Kajino, and Yoshiharu Ishikawa, "Algebraic Service Specification and Rule Generation for Integrating Multiple Dissemination-Based Information Sources", Proc. 7th International Conference on Database Systems for Advanced Applications (DASFAA 2001), Hong Kong, April 2001, pp. 344-351.
- [3] Yousuke Watanabe, Hiroyuki Kitagawa, and Yoshiharu Ishikawa, "Integration of Multiple Dissemination-Based Information Sources Using Source Data Arrival Properties", Proc. 2nd International Conference on Web Information Systems Engineering (WISE 2001), Kyoto, Japan, December 2001.
- [4] Zloof, M.M.: Query by Example, Proc. FIPS National Computer Conf., pp. 431-438(1975)