

アクティブデータベース上の拡張メディエータ言語

中田 優作 稲森 真二郎 小西 修

高知大理学部情報科学科

780 高知県高知市曙町 2-5-1

{yunakata,sinamori,konishi} @is.kochi-u.ac.jp

概要

メディエータは、利用者と情報源の仲介を行う情報システムである。特に、情報源は、半構造化データなどを含む異種情報源が扱われ、メディエータの役割として、スキーマ統合、情報統合が行われる。我々は、アクティブデータベース上のメディエーション・システム HI-AMS モデルを提唱している。本稿では、HI-AMS におけるメディエータと情報源の間のインタフェース言語について考察する。オブジェクト定義、ライフサイクル、ECA ルール、トランザクションの要素からなる情報アクセスエージェントの表現と、その表現に基づく、メディエータと情報源の間の知識交換、及びオブジェクトデータモデルの拡張について述べる。また、インタフェース・プロトコルとして義務論理が導入される。

キーワード: メディエータ、エージェント、義務論理、ECA ルール、インタフェース言語

An Extended Mediator Language on Top of Active Database Systems

Yusaku NAKATA Sinjirou INAMORI Osamu KONISHI

Dept.of Information Science, Faculty of Science, Kochi University

2-5-1 Akebono-cho Kochi 780 Japan

{yunakata,sinamori,konishi} @is.kochi-u.ac.jp

Abstract

The mediator is an information system which mediates between an enduser and information sources. The role of mediator is the exchange of schema or the integration of data for heterogeneous information sources including semi-structured data. We have been proposed an active mediation system: HI-AMS model on top of active database systems. This paper presents a mediator interface language between a mediator and sources. We show the expression of an information access agent consisting of the elements of object definitions, methods sequence, ECA rules, and transactions and based on the notion, the exchange of knowledge between the mediator and sources, and the extension of object schema. We also introduce the deontic logic as the interface protocol between a mediator and sources.

Key words: Mediator, Agents, Deontic logic, ECA rules, Interface language

1 はじめに

WWW(World Wide Web)によって実現された、分散された情報源という新しい環境は、従来の分散データベース技術での多くの概念の再考を促している。その中でも、非構造型を含む異種・分散された情報源を変換し、統合するという問題は重要である。また、WWWの非公式性と分散制御された特性という新しい環境は、異種的情報を統合するためのツールの重要性を指摘している。

我々は、この問題を解決する一つの方法として、エージェント型メディアータ・モデルを提唱し、そのプロトタイプ・システム・アクティブメディアーションシステム・HI-AMS(High Intelligent-Active Mediation System)を開発している。[5]その中で、システム機能の中核部分をなす、異種情報源の情報の収集と統合を行なう情報アクセスエージェントを拡張メディアータ言語によって表現する研究を行なっている。

関連研究として、メディアータによる異種情報源の統合(スキーマ統合を含む)に関する研究には、主として二つのアプローチがある。

一つは、メディアータと情報源のインタフェースを知識ベースで表現するもので、代表的な言語としてKIF/KQMLがあるが、非常に重たいインタフェース定義言語となっている。

二つは、異種情報源の環境をデータベースととらえて、グローバルスキーマや演繹データベースからのアプローチをとるものである。共通インタフェース定義語として、特に半構造化データに適用できるOEM[4]があるが、単純なオブジェクトデータモデルとしてしか表現できないため演繹データベースとして論理を導入する研究がある。[1]また最近、標準メディアータ言語として、OQLをベースにした言語が提案されているが、その中にルールを導入してより表現を高める必要性が指摘されている。[2]

我々のメディアータHI-AMSと情報源の関係は、オープン系のコンポーネント間のインタフェースとしてとらえられる。そこには単純なオブジェクトデータモデルによるデータ変換をベースに、そこにおけるデータと操作の一貫性制約をECAルールによって表現することによって、メディアータと情報源の関係を豊かで信頼性のあるものにする

ことができる。これは、第三のアプローチであり、メディアータ(異種情報源の統合)という目標を、アクティブデータベース上で実現するという研究はまだない。

HI-AMSは、適切なオブジェクトモデル(IDL)[6]インタフェースがアクティブデータベース上に定義されたものと見なすことができる。これは、CORBAのような開放系システム上の他の情報源と交換、協調できることを意味する。

つまり、HI-AMSの設計思想は、より高度なエージェント機能を有するアクティブデータベースを分散コンピューティング環境(基盤)に統合することであり、そこでは、HI-AMSは、分散されたアプリケーションが相互に協調し合う上での、一つのコンポーネントとして扱われる。

本稿では、まず、2章でメディアータの基本モデルについて述べ、つづいて3章では、メディアータのための定義語三層プロトコルについて、4章では、メディアータ情報源インタフェースについて説明する。5章は、本稿のまとめと今後の研究課題を述べている。

2 メディアータの基本モデル

2.1 概要

webによって実現された、分散された情報源を統合するための可能なアプローチとして位置づけられるメディアータは、ユーザの要求に見合って、種々の情報源にアクセスし、情報を統合する。その統合した大量のデータの集まりから有用な情報を取り出し、推定したり、抽象化したりする処理を行ない、ユーザに有益な情報を提供する。このメディアータは、ユーザと情報源の間の調整を行なう知的なエージェントであり、アプリケーションのデータベースなど情報源からの独立を保証し、データ/知識変換を行なう高位アプリケーションモジュールである。また、特定のサービスを提供するのに、複数の自律的なメディアータをネットワーク上に走らせることも可能である。この目標を実現するために解決されねばならない重要な問題は次のようである。

1. 半構造化データも含めた異種情報源の情報統合

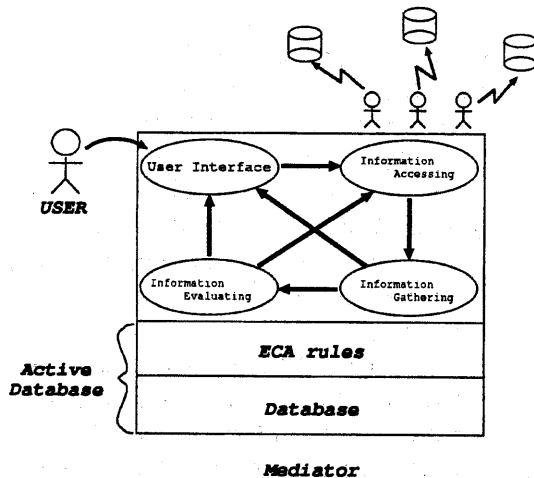


図 1: HI-AMS の構成

2. ユーザと情報源の間の調整を行なう知的エージェント機能
3. アプリケーションの情報源からの独立を保証するアーキテクチャ

これらの問題を解決するため、HI-AMS は次のような設計方針に基づいて開発されている。

1. エージェントの定義と制御を ECA ルールによって表現したアクティブデータベースを採用する。
2. 情報源やドメイン情報を自動収集するロボット機能を有する。
3. 異種情報源の情報統合のために、オブジェクトデータモデルと ORB の思想を取り入れる。
4. 情報融合と評価の機能の一つとして自己組織化処理機能を備える。
5. メディエータ同士は、コントラクトネットプロトコルをさらに拡張した義務論理によるプロトコルによって協調する。
6. web 環境のもとで利用するために、メディエータは web サーバと連携した 3 層モデルアーキテクチャを採用している。

2.1.1 情報エージェント

ユーザ・インタフェース・エージェント (質問・検索法) と情報アクセスエージェント (アクセス・集約)、及び情報統合・評価エージェント (情報解析・学習・探索戦略) からなり、ドメインや情報源の情報を収集し、複数の異種情報源にアクセスするロボット機能及び情報を集約し、自己組織化を行なう情報評価機能を有する。

2.1.2 アクティブデータベースシステム

情報エージェントはアクティブデータベースをプラットフォームとし、その ECA ルール機構によって制御される。これによって、柔軟かで、かつ信頼性のあるメディエータが実現される。以下に、我々のモデルの特徴であるアクティブデータベース上のメディエータの理由を述べる。

1. このメディエータ (エージェント) とその中の作業エージェントは、自分の行動に必要な情報や共通データを頻繁に使用するため、データ・情報を正確に管理できるデータベース機能が必要である。
2. エージェントの行動は、データベース上のアプリケーション (トランザクション) としてとらえることができ、エージェントの信頼性を高めるためにも必要である。
3. メディエータ機構のエージェントの仕事と作業はプロセスモデルでとらえることができ、そのプロアクティブなプロセス制御に ECA ルールを適用することは理にかなっている。
4. メディエータ機構は、アクティブデータベースの発展形でもある。

2.2 情報アクセスエージェント

情報アクセスエージェントは、我々の提唱する HI-AMS において、システムの中核をなす、異種情報源の情報の収集と統合を行なうエージェントである。以下にその機能を示す。

1. 質問文の受理
2. ソースへのスキーマの問い合わせ

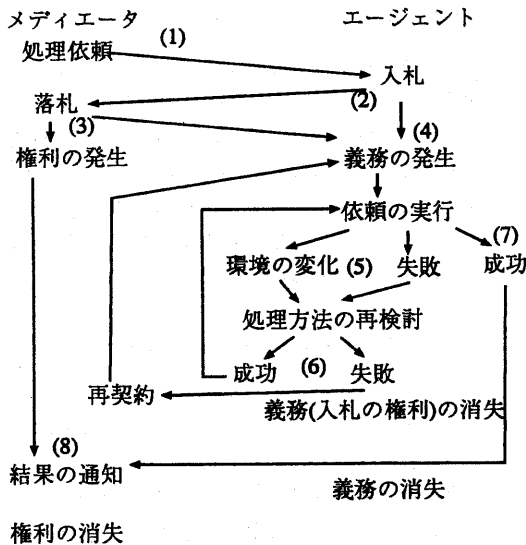


図 2: 義務論理プロトコル

3. ソースへの質問
4. 質問結果の受理
5. メディエータスキーマへの変換
6. 情報の統合

3 メディエータのための義務論理プロトコル

契約ネットプロトコル [7] を使用することにより、メディエータは支配下にある複数のエージェントに対して入札を行わせる。それによって、最適なエージェントを選び出し、処理の依頼を行うことができるが、そのエージェントを直接支配しているわけではないので、エージェントからの処理結果までを保証することはできない。エージェントが処理依頼前の状況から変化することや処理依頼に対する見積もり評価の誤りにより契約履行が不可能な場合が充分考えられる。そこで、メディエータとその支配下にあるエージェントとの関係に義務論理を導入する。従来の契約ネットプロトコルでは、落札終了後はエージェントからの実行結果待ちとなるだけであるが、拡張された契約

ネットプロトコル [8] では、エージェント側には入札したことに対する義務が発生し、メディエータ側には落札したことで権利が発生される。

そのため、エージェントは契約内容を処理し終えるまで義務を負うこととなる。しかし、エージェント内の状況の変化などにより入札時の方法では処理できない場合は、(1) 処理方法の再検討をエージェントの権限内で行う、(2) 再検討により処理不可能であるならば、義務の放棄をメディエータに告知する。

一方、権利の発生したメディエータは、義務を果たしたエージェントからの回答を受け取ることができる。しかし、義務の放棄をエージェントから通知を受け取ると、そのエージェントを一時的に支配下エージェントのリストからはずして、再度契約をやり直す。

このようにして義務論理プロトコルにおいては、実行主体を把握でき、環境の変化への適用力を増すことにより、分散された環境での運用性が向上する。

メディエータの動作の手順を以下に示す。(図 2 参照)

- (1) 依頼メッセージを配下のエージェント (複数) に送る。
- (2) メッセージを受け取ったエージェントは依頼内容を評価し、その結果を見積りとして入札する。
- (3) メディエータに入札された結果より、最適なエージェントを落札し、結果を通知する。落札した入札内容に対する権利が発生する。
- (4) 落札されたエージェントは、入札内容を実行すると共に、入札内容に対する義務が生じる。
- (5) 入札時から実行を終えるまでに情報源の環境が変化したり、入札の見積り評価に誤りがあったり失敗すると、エージェント自身で処理方法の再検討を行なう。
- (6) 処理方法の再検討に成功すれば依頼を実行し、失敗すればメディエータに通知して再び契約行動を起こす。(失敗したエージェントは、義務の消失とともに入札出来なくなる (禁止))

(7) 依頼された実行に成功すれば、実行結果をメディアータに伝え、入札に対する義務が消失する。

(8) 依頼した処理結果を受け取ったことで、メディアータの持つ権利も消失する。

4 メディアータ-情報源インタフェース

4.1 概要

ここで、メディアータスキーマを定義・表現するメディアータ言語が必要であるが、半構造化データなどの異種情報源を扱うため、共通プロトコル(データ変換のための共通インタフェース言語)として、オブジェクトデータモデル(TS IMMISのOEMやCORBAのIDLのような)を採用する。しかし、このオブジェクトデータモデルでは、オブジェクトの多くの意味制約や複雑なオブジェクトを扱うことはできない。これらの問題は、メディアータと情報源の間の情報や知識の交換によって克服することが考えられる。我々は、これをメディアータにおけるルールの使用によって解決する。それは、メディアータと情報源の間の関係を豊かなものにすることができる。

4.2 ルール表現

メディアータと情報源の間のインタフェース言語にアクティブデータベース上のECAルール機構を導入することにより、メディアータと情報源の間の一貫性制約を維持できる。

1. ルールによって、指定した情報源の内容上の一貫性制御を表現できる。
2. ルールは、質問処理とスキーマ情報の交換のために使用する。
3. メディアータと情報源の間のデータやスキーマ変換をルール形式で行なう。
4. ルールによって、情報源からの情報をメディアータのスキーマに統合する方法を指定する。

5. 情報源における更新がメディアータの質問に対してどのように反映すべきかをルールによって記述する。

6. ルールによって、情報源の更新に対するモニタリングを行なう。

7. メディアータへ渡された質問メッセージや、メディアータから情報源への質問がルールとして表現される。

HI-AMSでは、これらのルール機構は、アクティブデータベース上のECA機構によって表現される。そこでは、時間やデータベース操作やエージェントのイベントにより、より柔軟で、高度なルール機能を実現される。そのことによって、メディアータの一貫性制約が強化され、また、エージェンシー(推論能力)が増すことができる。

4.3 構成

情報アクセスエージェントは、インタフェース言語によって、オブジェクト定義、ライフサイクル、ECAルール、トランザクションという4つの要素で表現される。[3]ここでは、我々が考えるインタフェース言語での記述例を示す。

```
class ACCESS
  attributes:
    access_No : OID;
    num_hits : INT;
    num_data : INT;
    keyword : KEYWORD;
    source : SOURCE;
    receivedData[100] : DATA;
  methods:
    search(k:KEYWORD, s:SOURCE)
      { keyword:=k; source:=s };
    returnNumber(n:INT)
      { num_hits:=n };
    takeResult()
      { source.sendData(keyword) };
    receiveResult(d:DATA)
      { num_data:=num_data+1;
        receivedData[num_data]:=d };
  lifecycle:
```

```

search,returnNumber,takeResult,
receiveResult
rules:
staticRule:
    on before returnNumber(n)
    if n <= 0
    do abort
dynamicRule:
    on after receiveResult(d)
    if num_data != num_hits
    do abort
transactions:
    search, returnNumber, takeResult,
    receiveResult

```

これは文献情報検索におけるメディエータと情報源とのやりとりを記述した簡単な例である。情報源にキーワードにヒットする文献が何件あるのか質問し、情報を取得する。ただし、このやりとりには制約が定義されており、例えば情報源にキーワードにヒットする文献が一つもない場合には、アボートする。また、仮に文献があっても、最初の質問で受け取った件数と実際の情報の数が一致しない場合もアボートする。次に文献情報のための拡張オブジェクトモデルスキーマの簡単な例を示す。

```

class DOCUMENT
attributes:
    document_No : INT;
    title : STRING;
    authors : STRING;
    abstract : STRING;
methods:
    updateDocNo(n:INT)
        { document:=n };
rules:
    cancelUpdate:
        on after update(n)
        if exists d in DOCUMENT :
            (d!=self) and
            (d.document_No=self.document_No)
        do abort
transactions:
    updateDocNo

```

ドキュメント番号を変更しようとする時、既にその番号が使われている場合には重複してしまう。このような制約を取らないようにルールで記述する。このように、従来のオブジェクトデータモデルにルールを拡張することで、よりメディエータと情報源の間の関係を豊かで、信頼性のあるものにする事ができる。

5 おわりに

本稿では、アクティブデータベース上のメディエーション・システムと情報源の間のインタフェース言語について考察し、(1)義務論的プロトコルの有用性と、(2)ルール機構によって拡張することにより、メディエータの一貫性制約を強化できることを示した。今後の課題として、メディエータのエージェンシーを増すための、より柔軟で高度なルール機能を定義することである。

参考文献

- [1] Bergamaschi, S., "Extraction of Informations from Highly Heterogeneous source of textual Data", pp.42-63, CIA'97, LNAI1202, 1997.
- [2] Buneman, P., Raschid, L. and Ullman, L., "Mediator Language - a Proposal for a Standard", SIGMOD RECORD, Vol.26, No.1, pp.39-46, March 1997.
- [3] Gerti Kappel and Michael Schreffl, "Modeling Object Behavior: To Use Methods or Rules or Both?", LNCS1134, DEXA'96.
- [4] Hammer, J., Garcia-Mokina, H., Ireland, K., Papakonstantinou, Y., Ullman, J., and Widom, J., "Information Translation, Mediation, and Mosaic-Based Browsing in the TSIM-MIS System", SIGMOD 95.
- [5] 小西 修, "異種情報源統合のためのアクティブメディエーション・システム-HI-AMS:High Intelligent-Active Mediation system-", Mem.Fac.Sci.Kochi Univ. (Inform.Sci), 17, March 1997.
- [6] Object Management Group, "Common Object Services Specification Volume I.", OMG Document No.94-1-1
- [7] R. G. Smith, "The Contract Net Protocol:High-Level Communication and Control in a Distributed problem Solver", IEEE TRS. on Computers, Vol. C-29, No.12, pp. 1104-1113, 1980.
- [8] H. Weigrand, et. al., "Interoperable Transactions in Business Models - A Structured Approach", Advanced Information Systems Engineering, LNCS 1080, pp. 193-209, 1996.