

中田 優作 小西 修
高知大理学部情報科学科

WWWによって実現された非構造型を含む異種分散情報源を変換し統合するという問題に対して、我々は一つの解決方法としてエージェント型メディエータモデルを提唱し、そのプロトタイプ・システムとしてHI-AMS(High Intelligent - Active Mediation System)を開発している。[5] ここで情報収集と統合のため、メディエータと情報源の間の共通インタフェース言語が必要である。そこで本研究では、オブジェクトデータモデルを核言語とし、メディエータと情報源の間の一貫性制約を強化するためにECAルールによって拡張したインタフェース言語を開発し、異種情報源の統合を目的とする。本稿では、その概要について述べる。

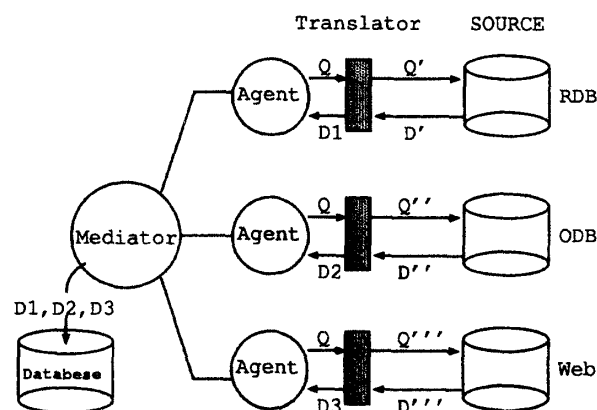


図 1: メディエータと情報源の関係

図1にメディエータと情報源の関係を示す。ユーザからの要求を受け取ったメディエータは、それぞれの情報源に適したエージェントを発行する。この時、メディエータはどの情報源にどのエージェントが適格なのかを知っているものとする。エージェントは各情報源に対してスキーマ情報を問い合わせ、メディエータからの要求を各情報源への質問文に変換して情報源に質問する。情報源からの質問結果はメディエータスキーマに変換され、統合される。

TSIMMIS の OEM や CORBA の IDL のようなオブジェクトデータモデルが適していると考えられる。オブジェクトデータモデルにおけるインタフェース定義では、オブジェクトの性質や状態を表す属性や関連といったプロパティとオペレーションを定義する。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ができない。そこで、このオブジェクトデータモデルを核言語とし、ECA ルールにより拡張されたメタ・インタフェース言語を考える。つまり基本的には、軽いオブジェクトデータモデルとその一貫性制約を強化する ECA ルールからなるという考え方に基づいている。

3.1 概要

半構造化データを含む情報源を変換し統合するため、メタデータと情報源の間の共通インタフェース言語が必要である。その共通のデータモデルとして、

3.2 情報エージェント

情報収集を行なうエージェント、すなわちメディアータと情報源の間のやりとりをインタフェース定義言語を用いて表現することができる。情報エージェントは agent (attributes, methods, lifecycle, rules, transaction) という構成で表現する。[3] この時ルールは、メディアータと情報源の間の情報の伝達用言語としての役割を持つ。

3.3 メディエータスキーマ

情報を統合するためのメディエータスキーマをインタフェース定義言語を用いて表現する。図2に文献情報を扱うためのメディエータスキーマの簡単な記述例を示す。

```
interface Document
(extent Documents){
  attribute string docNo;
  attribute string title;
  relationship Set<Person>
    authors inverse Person :: writes;
  void updateDocNo(string n);
  trigger cancelUpdate :
    on after updateDocNo(n), create()
    if exists d in Document :
      (d!=self) and (d.docNo=self.docNo)
    do abort;
};
interface Journal : Document
(extent Journals){
  attribute string abstract;
};
```

図2: メディエータスキーマの例

型のインタフェース定義はキーワード `interface` によって始まる。上位型を持つ場合は型名の直後に上位型名を記述する。この例の場合、`Document` 型は `Journal` 型の上位型であり、`Journal` 型は `Document` 型を継承する。`attribute` で始まるのが属性の定義であり、ドキュメント番号とタイトルは基本データ型 `string` のデータである。また、`relationship` 以下に型の関連を定義しており、著者は `inverse` によって相互参照が可能な集合型のデータとして記述されている。これらのデータを操作するためにメソッドが用意されており、ここではドキュメント番号を更新するためのメソッドとその引数を記述してある。ここまでの通常のオブジェクトデータモデルで表現されるインタフェース定義であるが、メディエータ・インタフェース言語では、ルールによって指定した情報源の内容上の一貫性制約を表現する。この例では、`Document` 型とそれを継承する型のオブジェクトを生成、またはそのドキュメント番号を変更しようとした時に、既に存在するドキュメント番号との重複を許さないという制約を守るためのルールを記述してある。このように、ルールの形での知識の伝達を行なうことによって、メディエータと情報源間の関係をより豊かなものにすることが可能である。

3.4 統合

メディエータスキーマに変換された情報源は、メディエータのプラットフォームであるアクティブデータベース上で扱うことができないといけない。そのため、ルールも OQL のようなアクティブデータベース上の質問言語で処理されることが必要である。

4 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の開発についての概要を示し、インタフェース言語を ECA ルールで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、Java - C++ - Ode といったシステムによって実装する予定である。

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 Schrefl, "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 TSIMMIS 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