

異種分散環境におけるエージェント指向型情報検索システム —設計方針と基本構成—

7P-1

箱守 聡 横山 和俊 田辺 雅則 井上 潮

NTTデータ通信（株）情報科学研究所

1. はじめに

近年、様々な情報が電子化されており、これらを格納したデータベースをインターネットに代表されるコンピュータネットワークに接続することによって、大規模な異種分散環境が構築されつつある。このような環境では、利用者が求める情報を効率的に検索できるシステムの提供が求められている。

本稿では、多種のデータベースが存在する異種分散環境における情報検索の問題点と要求条件を明らかにし、これらを満たすエージェント指向の情報検索システムを提案する。システムの設計方針と基本構成、および実現上の課題について述べる。

2. 情報検索に対する問題点と要求

多種のデータベースが存在する異種分散環境において、利用者が情報検索を行なおうとする場合、以下の問題がある。第一の問題は、利用者が利用可能な全てのデータベースについて知ることが不可能であり、また各データベースの内容が随時更新されるため、求める情報がどのデータベースに格納されているかがわからないことである。第二の問題は、データベース毎にアクセスインタフェースが異なっているため、利用者が検索時にそれらの相違を意識しなければならないことである。これらの問題により、利用者が必要とする情報を的確に得ることは困難であり、また情報を得るために大きな労力が必要となっている。このため、必要な情報を簡易に取得できる情報検索システムの実現が求められる。システムに対する要求条件を以下に示す。

(1) データベースの選択

利用者の要求に対して適切なデータベースを選択することが求められる。利用者の要求には、情報の内容に対する条件と、処理時間や利用料金等のシ

テム利用に対する条件とがあり、これらを満たしたデータベースを選択する必要がある。

(2) 検索処理の代行

選択したデータベースに対してアクセスを行ない、得られた結果を統合して利用者へ提示することが求められる。このとき、一つのデータベースを検索した結果得られた件数が少ない場合にはさらに別のデータベースへアクセスする等、処理結果が利用者の要求を満たすようにすることも必要である。

分散環境における情報検索システムの研究には、Lycos[1], WAIS[2], GIOSS[3]等がある。LycosはURL(Uniform Resource Location)を対象とし、WAISは独自のデータベースを使用するため、どちらも既存のデータベースが持つ情報を利用できない。また、GIOSSは検索処理の代行を行なわないため、利用者が直接データベースへアクセスしなければならない。

3. 情報検索システムの提案

3.1 設計方針

2章で述べた要求条件を満たす情報検索システムを提案する。システムの設計方針を以下に述べる。

(1) 分散型のシステム構成

情報検索システムには、集中型と分散型の2通りのシステム構成が考えられる。集中型では統一したアクセスインタフェースを提供できるが、膨大な量のデータを蓄積し操作するための高性能な計算機が必要である。一方、分散型は既存のデータベースを利用するため、効率的である。このため、分散型のシステム構成とする。

(2) エージェント指向

システムの環境は頻繁に変化しており、これらの変化に応じて動的に処理対象や処理手順を決定する必要がある。このように柔軟なシステムを構築するために、エージェント指向の概念を導入する。エージェントの持つ自律性や協調性等の性質を用いることにより、データベースの異種性の隠蔽やシステムの状態に応じた動的なデータベース選択や検索処理の実行が実現できる。

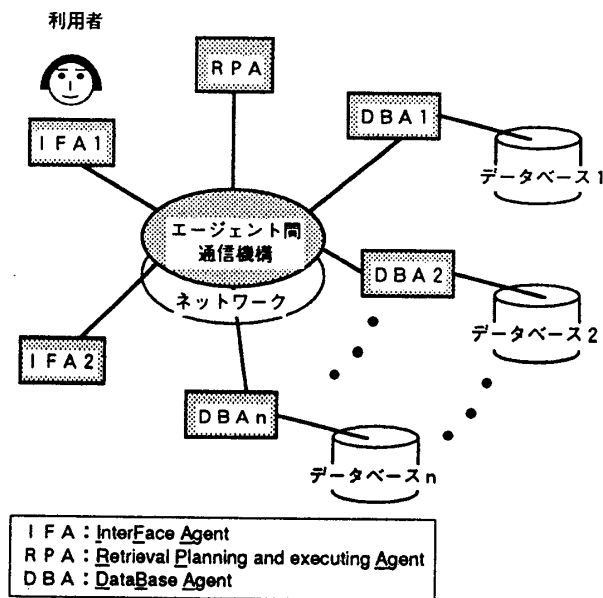


図1. 情報検索システムの構成

3. 2 システム構成

提案する情報検索システムの構成を、図1に示す。システムは、データベース、3種類のエージェントおよびエージェント間通信機構から構成される。各エージェントの機能を以下に示す。

(1) インタフェースエージェント(IFA)

利用者とのインタフェースを受け持つ部分であり、利用者との対話を行ないながら、検索条件式や取得件数の上限等からなる情報検索の要求メッセージを作成する。また、検索結果を利用者に分かりやすい形で提示する。

(2) 情報検索代行エージェント(RPA)

情報検索の要求メッセージに従い、利用者の要求に対して最も適切なデータベースを選択する。このため、各データベースの内容や利用料金等の情報を保持する。また、選択されたデータベースに対して情報検索の実行を依頼し、得られた結果を統合してIFAに返却する。

(3) データベースエージェント(DBA)

各データベース毎に存在し、データベースへのアクセスインタフェースを受け持つ。RPAからの依頼を受け、それぞれのデータベースに最も適した方法で検索処理を実行し結果を返却する。また、データベースの稼働状況や利用料金等の情報をRPAに通知する。

また、エージェント間通信機構は、エージェント

間の機能やインタフェースの差異を吸収し、システム構成の変化に対応できる通信インタフェースを提供する。

4. システム実現上の課題

情報検索システムを実現する際の課題を以下に述べる。

(1) ユーザインタフェース

どのようにして利用者の意図を明確化し、正確に反映した要求メッセージを作成するか

(2) データベースの選択方式

データベースの特徴の表現方法やその取得方法、それらを用いたデータベース選択方法[4]

(3) 検索処理の実行方式

データベース間で異なるインタフェースの統一方式や、検索処理のスケジュール方法

(4) エージェント間通信

様々な種類のエージェントに対応する通信インタフェースの実現と、エージェントに関する情報の管理方式[5]

5. おわりに

多種のデータベースが存在する分散環境を対象としたエージェント指向の情報検索システムを提案した。現在、ワークステーション上でプロトタイプシステムが動作中であり、今後はこれを用いて各課題に対する実現方式の評価を行なっていく予定である。

参考文献

- [1] Mauldin, et al.: "Web Agent Related Research at the Center for Machine Translation", SIGINDR meeting(1994).
- [2] Kahle, et al.: "An Information System for Corporate Users: Wide Area Information Servers", Tech. Rept. TMC199, Thinking Machines Co.(1991).
- [3] Gravano, et al.: "The Effectiveness of GLOSS for the Text Database Discovery Problem", ACM SIGMOD, pp.126-137(1994).
- [4] 田辺他: "異種分散環境におけるエージェント指向型情報検索システム—データベース選択方式の評価—", 第52回情処全大(1996).
- [5] 横山他: "異種分散環境におけるエージェント指向型情報検索システム—CORBAを用いたエージェント間通信機構の実現—", 第52回情処全大(1996).