

地域データ利用のための Query-by-Object 型インターフェイス構築に関する検討

出水田智子¹, Subhash Bhalla², Nadia Berthouze²

¹会津大学コンピュータ理工学部文化研究センター

²会津大学コンピュータ理工学部コンピュータ・ソフトウェア学科

概要

都市や地域を対象とした研究分野では、国や地方自治体などが整備する地図や統計などの様々な地域データを利用することが多い。これまでは簡単なデータ検索や表示でも、独自のプログラムを作成するか、地理情報システム(GIS)などの専用のアプリケーションを使用するしかなかった。データの入手も、磁気テープやCDなどの物理的な記憶媒体を介して行われてきた。近年 Web 技術の向上と普及により、データの閲覧・入手・検索などが Web 上でも簡単にできるようになった。しかし複雑なクエリーを発行するためには SQL 文の作成を必要とするものがまだ主流であり、フォーム形式で簡単に SQL 文を作成できる場合もあるが、利用できる SQL の機能が限定されることが多い。特に空間的なデータベースの検索インターフェイスは機能が限られた定型的なものが多いため、ユーザの自由な視点からの情報取得が困難である。本稿では、プログラミングやデータベースの技術的な経験や知識が少なくても SQL と同等なレベルの情報検索を可能とするオブジェクトによる検索(Object-by-Object)方式のインターフェイスを提案し、地域データの共有と再利用を図るのための統合的なデータシステムに適用する方法を検討する。

Abstract

Recently, with the advent of Web based information technology, more detailed and large-scaled regional datasets such as geographical and statistical data are available for both research and general use. Data browsing, distribution, and enquiry are becoming easier through the Web. However, most query interfaces limit query objects and paths at the implementation level, which are fully supported by SQL for highly skillful experts. We propose a user interface to support SQL query based on object-based calculations termed as Object-by-Object Calculate (OBOC). The interface equips the skill less users with object level query language tools. Some advantages of the proposed interface applied for Web-based regional data sharing system were examined.

1. はじめに

web 上の情報システムの到来とともに、高度な言語をユーザが直接やりとりできるように支援することが必要となってきている。ユーザの自由なやり取りを保障するためにはDBMSの高度なクエリー言語である SQL を表現できるようにしなければならない。われわれが提案する Query-by-Object 方式の検索システム([1])では、WWWを通してサポートされているDBMSア

アプリケーション上でユーザが高度な言語をやりとりできるようになっている。提案された言語における段階的な進め方は検索対象と検索経路を論理的に追っていくだけでなく、リレーショナルデータベース管理システム(RDBMS)により提供されている SQL をサポートされている。提案された QBOCアプローチは直感的である。また、関係代数の演算がサポートされているため、関係性が保たれている言語である。このインターフェイスの利点はプログラミング技能の既得をユーザに要求しないことである。ユーザは自然で論理的なアプローチを用いることによってテーブル閲覧形式のナビゲーションを手順を踏みながら実行できる。また複雑なクエリーの実行をサポートし、クエリーがシステム内部でどのように操作されているか確認しやすくしている。なお、試行中のシステムの基本的な概念や利点については文献[1]で詳しく説明している。O-RDBMS を含む空間データベースの概念やシステムについては文献[8]が詳しい。ここでは簡単にこのインターフェイスの主な特徴を列挙する。

[O-RDBMS の検索インターフェイスとしての特徴]

- ・既存のシステムでは、コマンドラインによる SQL 文の直接の入力を行うことによって利用することが可能であったリレーショナルデータベースの高度な検索を可能にしている。
- ・既存のシステムでは用途に合わせてた定型処理目的のダイアログによるいわゆる QBE 方式の検索が多いが、そのようなシステムのもつ限界を克服した。
- ・空間データを扱うことが可能な DBMS に対応しているため、値や項目による属性検索だけでなく、特に地域データベースで必要とされる空間検索にも機能を拡張することができる。
- ・段階的な探索を行うことができるため、時間はかかるが、ユーザの自然で直感的な検索活動に沿ったインターフェイスを提供している。

[Web システムとしての特徴]

- ・更新されるデータの場合、ユーザは最新のデータにアクセスしやすい。
- ・分散したリソースをネットワークを介して一元的に管理できる
- ・エンドユーザの環境を選ばない（ブラウザソフトは必要）
- ・Web 技術や Web リソースを活用できる
- ・メタデータによる所在の確認からデータの入手までのユーザの一連の行動をサポートできる

2. 地域データの検索方法

3. 1 オブジェクトによる SQL 検索

現在われわれの研究グループが試作している Web 上で動作する検索インターフェイス（仮に WebQB0 と呼ぶ）は Web サーバとクライアントで構成されている。SQL で表現される命令文を表形式のダイアログを使用して作成し、Web を介してデータベースにアクセスするためのインターフェイスである。主な機能は、データをオブジェクトとして操作の対象化した関係代数と集合演算に基づくデータ集合の計算である。したがって検索結果はユーザが利用しやすいテーブルの形式で与えられる。データベースに格納されているデータは、リレーショナルデータベースとして管理

されているが、拡張した機能としてデータと手続きをひとまとめにしたオブジェクトの機能を扱うことができる。また、空間データを扱うことも可能なため、将来的には空間検索エンジンとしての拡張も可能であり、GIS や Web データベースなど、応用の範囲は広い。Web-GIS インターフェイスとしての過去の応用例は文献[2]で紹介しているので参照されたい。

3. 2 WebQBO の概要

サーバ側では Java 言語(1.5.0)で記述したプログラムがサーブレットエンジン Tomcat Servlet/JSP (5.0.27)で作成した Web アプリケーションを介して動作する。QBO の主要な機能は、オブジェクト指向に拡張されたリレーショナルデータベース管理システムである PostgreSQL (7.4.7)の採用によって実現されている。クエリーの結果を XML に出力するために JDBC (7.4.215)のドライバを使用している。クライアント側は、結果の出力を閲覧するために上記のサーバ環境に対応する Web ブラウザを必要とする。検索の実行は対象となるデータ集合のオブジェクトとオペレーションの組み合わせの選択および条件の入力によって行われる。1つ1つの演算が独立しているため、検索順序に依存しない、絞込み検索に有効な段階的検索アプローチを採用している。利用できる機能は SQL による命令文によるものとほぼ同じである。

地域データの利用を前提とした場合には、既存の GIS アプリケーションのもつ検索機能と同等な機能が要求される。空間データの利用にはさまざまな障害が残っているが、単なる情報のインデックスや文字や数値のデータ検索だけでなく、地図や画像や図形情報も利用できるマルチメディアな機能が求められる。なお、汎用的な GIS では3種類の検索が用意されている。

[空間検索]

地図上での対象図形の直接の選択や、矩形や円などの図形の輪郭線の重ね合わせを用いた選択、既存図形の選択による内部図形の選択など、主に対照となるデータの選択・表示のためのものと、空間データに基づいて距離演算や集合演算的な操作を行う計算的なものとに大きく分かれる。

[属性検索]

大きくは2種類の方法が提供されている。一つは一定の条件に基づく論理的な操作によって直接対象を選択する方法、もう一つはクエリーによって抽出する方法である。多くの GIS アプリケーションでは、選択あるいは抽出に用いた条件やその結果を保存して再利用することが可能となっている([8])。

PostgreSQL は、構造的な空間データもオブジェクトとして管理できるように、RDBMS に OODBMS の特徴を組み込んで拡張した O-RDBMS と呼ばれるデータベース管理システムである。このようなシステムは、表形式のデータ集合の一部として階層や複雑な構造をもった様々な型のデータを一元的に取り扱うことが出来るため、統合化データベースの構築に適している([3][7])。

4. 地域データシステムへの応用

4. 1 統合的な地域データ利用システムの提案

ここでは例として、地域的な調査研究の基本的なデータである国や自治体が提供する地図や統計の電子的なデータの作成・配布・利用を統合的に実現するための統合型データベースの検索エンジンとしてQBO方式の導入を検討する。上記の3つの機能はそれぞれ統合的なデータベースを構成する、基幹データベース、データウェアハウス、クリアリングハウスにおいてサポートされ、有機的に連携することによって威力を発揮する。そのフロントエンドの部分をQBO検索インターフェイスで構築することにより、期待される効果と実装時の問題点を以下に簡単にまとめる。ここで考えるデータウェアハウス、クリアリングハウスと類似のシステムとしては、東京大学空間情報科学センター(CSIS)の統計データベースシステムと空間データクリアリングハウスがある。

4. 2 基幹データベースにおける応用

基幹データベースとは、データベースシステム全体の核となる部分を構成するデータベースであり、実体のモデル化やデータの生成・更新・追加・削除など主にデータ蓄積機能を担う。標準的なSQLではDBMSを管理するために、これらの操作を実行させることができる。このデータベースのユーザは、対象とするDBMSを直接管理する権限を持つ人や、データの作成を行う人である。QBOでは上記の機能をオブジェクトとしてもつことによって、データ作業の効率化を図ることができる。また、視覚的なインターフェイスを実装することによって、SQLの知識が少ない人でも簡単にデータベースを作成することができる。また、Webで実装した場合、データの更新作業を遠隔地で行うことも可能になる。ただし通常のデータ作成作業は定型的な作業が多いため、作業の能率から見るとQBE方式のインターフェイスが優れている。したがってデータの生成に関わる基幹データベースに関してはQBOを採用するメリットはデータウェアハウスやクリアリングハウスと比較して少ないものの、システムの規模が小さいときは、移動時に基幹データベースから部分データを抽出する対面式の補助的なインターフェイスに応用できる。

4. 2 データウェアハウス

データウェアハウスとは、基幹データベースに時系列に蓄積されたデータを検索・分析して有効活用することに特化したデータベースを意味する。数年単位で継続的に実施される統計データファイルなどは基幹データベースで作成された後のデータ更新の問題が原則としてない。また統計データを行政の意思決定や住民との情報共有化に役立てるためには、時間情報を含む大規模な多次元データの分析や要約を必要とする。このため統計データベースは基幹データベースをソースとするデータウェアハウスとして分けて構築することが望ましい。

データウェアハウスでは検索や分析の目的に応じたOLAP(On-line Analytical Processing)を組み合わせることで高度なデータ利用が可能となる。またWeb上で機能させることにより、このようなシステムの利用がユーザサイドで可能となる。地域データの多くは種類や形式が多種多様で大規模なために、一定の条件で部分的に切り出されたデータファイルとして、分割提供されている場合が多い。また多くの統計データは、地域や主題別にポータブルなCDなどの記憶媒体やWebサイトを通じて利用者に配布されている。このようなデータの切り出しを多様でしかも変化するニーズをもつユーザ自身の手で行うことが可能になれば、データの利用に供するだけでなく、データ流通を行う国や自治体の業務やコストの軽減が期待できる。Web上で大規模データの参照

や検索を行うには容量や処理速度の問題を神田獲なければならないが検索によってデータを小分けできるので、ダウンロードできるサイズに制約を設けることで、サーバや回線に大きな負荷をかける危険を避けることができる。

リレーショナルデータベースに対して柔軟な検索機能を持つ QBO の検索インターフェイスはある種の OLAP であるとみなせる。そこで、OLAP として多様な条件検索でデータの切り出しやレポート作成を行う機能を付加すれば、QBO は統計データ発行システムの窓口インターフェイスとして利用できるかもしれない。さらに強力な多次元 OLAP を組み込むことによって、データマイニングの機能をもたせれば利用用途の拡大につながる。

4. 3 クリアリングハウス

様々な機関や個人が提供する地域データを自分の目的や必要にあわせて利用するためには、入手可能な各データに関する説明情報、いわゆるメタデータのデータベースがあると便利である。ある情報に関連するデータを収集・整理・配布するための情報検索システムをクリアリングハウスと呼ぶ。1994 年のアメリカ合衆国の大統領令で提案された国土地理空間データクリアリングではじめて使用された。このため、地図や統計などの地理情報を扱うものを指して使われる場合が多い。特に行政機関がそれぞれ作成・保有するデータのインデックスを管理する目的でメタデータのデータベースとなっている。このような社会的なシステムの運用によって社会的にも二重投資が避けられ利用が促進されるため、国内の地理情報については、「国土地理院地理情報クリアリングハウス」が 2000 年から Web 上で公開・運用されている。国内では空間データの作成者、時期、整備範囲、品質、入手方法などの項目が JMP: Japan Metadata Profile と呼ばれる規格で標準化されている。クリアリングハウスにはこの規格で記述された地理院のメタデータや各省庁・研究機関へのリンクが登録されている。メタデータの項目は多岐にわたるため、必要な情報を効率よく検索するためのインターフェイスが Web 上で公開されている。

研究者を対象とした空間データのクリアリングハウスは東京大学空間情報科学センターが 1998 年から Web 上で公開・運用している。空間データの座標系情報、緯度経度の数値、地名、など多様な条件検索や空間検索の機能を実装している。研究者のニーズ調査に基づくきめこまかい検索方法が準備されているものの、メニュー画面からできる選択肢は限定されている。分散リソースのリンクと検索という面では、異なる場所にあるデータリソースに対して、Web を介して検索する機能を付加できれば、クリアリングハウスの連携に役立てることができる。

4. 4 システムの課題

このインターフェイスの最も大きな特徴は、データを中心としたシステムの構築を行うことによって、プログラミングやデータベースの十分な知識がないユーザでも自由な視点でデータの検索・抽出・分析が可能になる点である。検索経路の自由な設定が可能とはいえ、集合演算や関係代数の考え方に沿った検索の段取りをイメージすることが必要となるため、その手段もサポートしなければならない。またこの方式は他の方式に比べてステップ数が多いために時間がかかる検索方法でもある。目的にあった QBE の検索方式は機能を制限することで使いやすくしているメリ

ットもあるため、QBOは用途に応じた検索インターフェイスの候補の一つとしてここでは考えている。

5. さいごに

本稿では、オブジェクト指向型の特徴を併せ持つRDBMSに対して高度な検索を行うためのWeb上のインターフェイスを提案し、地図データや統計データなどいわゆる地域データの作成・利用・配布の各段階に対応するデータベースでの応用と構築上の問題点を検討した。その結果、特にクリアリングハウスの構築に役立つ機能をWebQBOはもつことが明らかとなった。

今後システムを実装していくためには、既存の検索システムとの機能比較やニーズ分析に基づく必要な機能の追加など、検討しなければならない課題が数多くある。まずは地域データを入手する（提供側からの視点からは配布する）システムの機能を高めるためのインターフェイスをして、基本的な検索エンジンの実装を行っていきたいと考えている。

参考文献

- [1] A. Bhalla et. al (2005): An Interface for Web Based Access to Dynamic Contents. *Databases in Networked Information systems*, LNCS3433, Springer-Verlag, Berlin Heidelberg, pp. 70-77.
- [2] A. Bhalla et. al (2002): A User Interface for a Web-Geographic Information System (Web-GIS). *Databases in Proceedings of the 6th World Multiconference on Systems, Cybernetics, and Informatics*, vol. XVII, Orlando(Florida), pp. 217-222.
- [3] M. Schneider (1997): *Spatial Data Types for Database Systems*. LNCS1288, Springer-Verlag, Berlin.
- [4] P. Rigaux, M. Scholl, and A. Voisard (2002): *Spatial Databases with Application to GIS*. Academic Press, Morgan Kaufmann, San Francisco.
- [5] 岡部篤行ほか (2002): 『学術空間データ基盤システムの構築-東京大学空間情報科学研究センターの事例-』. Sinfornica 叢書, 財団法人統計情報研究センター, 東京.
- [6] CSIS クリアリングハウス <http://chouse.csis.u-tokyo.ac.jp/gcat/editQuery.do>
- [7] 高阪宏行・岡部篤行編(1996): 『GIS ソースブック-データ・ソフトウェア・応用事例-』. 古今書院, 東京.
- [8] Yue-Hong Chou(1997): *Exploring Spatial Analysis in Geographic Information Systems*. Onword Press, Santa Fe.