

観光情報の登録並びに検索を支援する 実験システムの設計と評価

永瀬 宏 久保澤 靖恵 渡辺 太一
金沢工業大学 人間・情報・経営系

〒921-8501 石川県石川郡野々市町扇が丘7-1

地域観光情報を提供する WWW ホームページの開設を想定して、観光施設による WWW 情報のコンテンツの登録、並びに一般ユーザによる登録済コンテンツの検索を支援する実験システムを試作した。本システムでは登録情報や検索条件の編集をクライアント側で行い、サーバ側のオブジェクト指向データベースへは簡素な手順でのアクセスを可能とした。コンテンツの検索ではユーザの興味に応じて複数の視点を設けている。視点には名前の集合からなる離散的視点と、類似度が定義される位相的視点があり、ユーザは各視点で任意に検索条件を指定する。これらの組み合わせはクライアント側で整合性を調べ、検索キーを組み立ててサーバに検索を依頼する。この結果、比較的に自由度の高い検索がサーバへの負荷をかけずに実現できている。

Design and evaluation of an experimental system for supporting registration and search of tourist information

Hiroshi Nagase Yasue Kubosawa Taichi Watanabe

Kanazawa Institute of Technology

7-1 Ougigaoka Nonoichi Ishikawa 921-8501 Japan

An experimental system is designed to offer an imaginary WWW home page that provides local tourist information. The system has capability to register contents of each WWW page, corresponding to tourist spot, and to search registered contents. Since editing registration items and search conditions is performed at client's sites, simple accesses to object oriented databases are accomplished at server's sites.

For reference of contents, several viewpoints are assumed according to users' interests. Viewpoints are classified to discrete views, composed of a set of names, and to topological views, where similarity measure is defined. Users specify their search conditions, under selected view. Client's sites examine consistency of specified conditions, construct search keys and send them to servers. Totally, proposed system realize a flexible navigation tool without imposing overloads to servers.

1.はじめに

本稿は地域内の観光地やイベントに関する WWW ページの URL やその内容をより迅速に検索できるようなシステムの構築を目指したものである⁽¹⁾。これは言わば地域コミュニティのナビゲータであり、技術的には WWW ページのコンテンツをサーバに蓄積することである。各観光地などの WWW ページの内容を 1 次情報とすれば、コンテンツは標準的な内容に絞った 2 次情報であり、コンテンツを探すためのキーが 3 次情報である。この 3 階層の情報構造の中で、特に 2 次情報を充実することにより、よりユーザの意図に近い情報を検索可能とするとともに、場合によっては 2 次情報だけでも十分に検索の目的を達成できる可能性がある。本研究はこのような意図に基づき、地域コミュニティで独自に WWW のナビゲータを作成する手助けをすること、その際、情報検索の平凡ではあるが自然な方法を実現することの 2 点にある。

地域コミュニティのナビゲータの構成には概ね、①汎用性のある大規模な商用システム⁽²⁾、②主として公共団体などが管理するホームページ⁽³⁾、③市販のナビゲータ構築ツール⁽⁴⁾の利用、の 3 通りの方法が考えられる。いずれも現在、利用されている手段だけに利便性の高いものであり、コンテンツの提示機能も簡単な解説を付加することで部分的に実現されている。しかし、登録情報更新の容易さ、検索画面の自由な設計、運用コストの低減等を考慮すると、独自に検索用の WWW ページを開設できることが望ましい。

ナビゲータにコンテンツを付与する場合、その作成方法や登録の仕方が直ちに問題になる。コンテンツの作成では①センター依存、②Web サイト依存、③自動収集の 3 通りの方法が考えられる。このうち、登録情報更新の容易さから Web サイトのユーザに依頼するか、ページの更新を自動検出することが望ましい。本研究では現時点で実現可能な②Web サイト依存の方法を採用した。これは Web サイトからの要求に基づき、2 次情報登録のための枠組みをサイトに提示して、そこに記入して返送してもらうという、言わばアンケート形式である。この方法によれば、サイトの WWW ページを自動解析して、予め枠内を埋めておき、サイト側のユーザに訂正してもらうことで、③自動収集の機能も部分的に実現可能である。但し、現在のところ WWW ページの自動解析はまだ行っていない。

以上に述べた方法は現在でも CGI により実現できるものである。しかしサーバ側で登録画面の編集や受信情報の分析が必要であり、負荷を増大させる恐れがある。そこで本研究では Java アプレットを用いて作業の大半をクライアント (Web サイト) に移管する方法を採用した。なお現時点での試作システムは単一サーバであるが、将来的にはサーバを分散させる計画であり、その中ではクライアント側でサーバを選択する機能を持たせる予定である。

登録情報の検索でも、その実現方法は登録時とほぼ同等であり、検索のための枠組みをユーザに提示して希望する項目を選択させている。ここでも Java を利用している。但し、コンテンツの検索ではユーザの興味に応じて複数の視点を設けていることが特徴となっている。視点には名前の集合からなる離散的視点と、類似度が定義される位相的視点がある。離散的視点の具体例は行動 (見る、買うなど)、対象 (音楽、美術など) 他、複数存在する。従来からのキーワード検索では検索項目の AND/OR 式をキーとするが、このような式を作成することは一般ユーザには専門的に過ぎる。離散的視点では項目間の関係をユーザが特に明示せず、コンテンツサーバの内部処理で対処する。位相的視点は距離、時間などであり、検索条件が連続量、あるいは定性的条件となる。ユーザは各視点で任意に検索条件を指定する。これらの組み合わせはクライアント側で整合性を調べ、検索キーを組み立ててサーバに検索を依頼する。この結果、比較的 to 自由度の高い検索がサーバへの負荷をかけずに実現できている。

本研究では検討内容に基づき、実験システムを試作した。サーバにはフリーのオブジェクト指向データベースである ObjectStore™を使用し、ソフトウェア全体が無償で提供できる内容となっている。

2. ユーザの視点に立脚したコンテンツの分類

2.1 人間の行動モデル

地域コミュニティの主として観光情報が登録されている、WWW のホームページを開くと仮定する。これは各 WWW ページのコンテンツ登録、並びに検索機能を合わせ持つ。このうち検索の部分、すなわちナビゲータを設計するに当たり、ユーザがいかなる目的でこれを利用するのかを予め、分析する必要がある。そこで多少、独善的かもしれないが、ユーザが情報検索を目的として本ページにアクセスするときの行動を図1のようなシナリオに描いた。

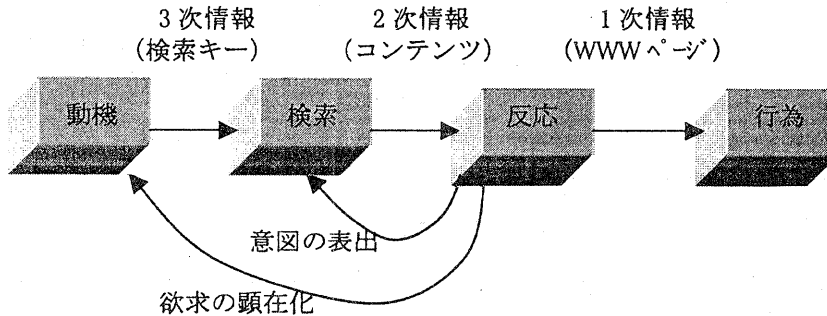


図1. 情報検索の行動モデル

人間が端末からアクセスするとき、そこに必ず動機が存在する。動機には人間に内在する本質的な欲求と、それを顕在化するための意図が存在すると考えた。このうち人間の欲求としては知識欲、所有欲などが先ず想定されるが、観光案内に限れば雑誌のキーワードの如く「見る、食べる、遊ぶ」がある意味で欲求の本質をついているし、現代社会においては「時間を潰したい」なども広義の欲求と考えられる。

欲求が人間に関わる属性であるのに対して、その欲求を達成するためには、機械に分かる言葉に翻訳して伝達する必要がある。その言葉をここでは意図と呼ぶ。意図の具体例は観光地の名前を調べる、コンサートのチケット予約をするなどである。ナビゲータ上で検索を行いながら、その結果を見て次の行動を選択する反応の過程は自身の意図を明確化する手続きである。これをここでは意図の表出と呼んでいる。ところで、ユーザによっては欲求が潜在化していて、自身が何をしたいのか明確でない場合も想定される。図中、欲求の顕在化と書いた、反応から動機へのパスはこのような状況を表現している。

これらの過程を経て、然るべきコンテンツ (2次情報) が得られ、当該地に出掛ける等の具体的な行為に移っていく。

2.2 意図の表出

ユーザに検索の目的を明示させ (意図の表出)、結果を提示する手続きにおいて、ナビゲータは①検索メニューの提示、②検索キーの整合性検査、③検索ならびに結果の表示からなる3段階の処理を行う。

(1) 検索メニューの提示

情報検索のためのメニュー提示では通常、多分木構造の選択子が用いられるが、メニュー項目の選択手順がシステム設計時に固定されてしまう問題点がある。そこで本試作では最初に述べたように、選択子を複数の視点にグループ分けして、複数の視点について同時に項目指定できる検索方法を採用した。実際のメニュー画面の例を図2に示す。

視点には名前の集合からなる離散的視点と、類似度が定義される位相的視点がある

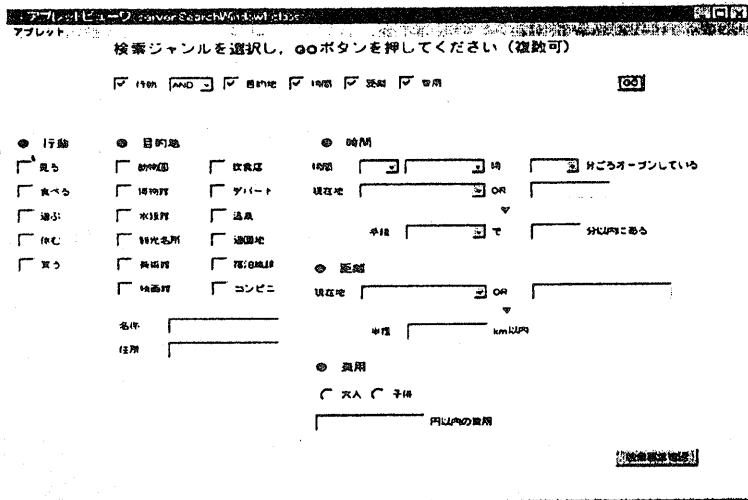


図2 検索画面例

考えられる。離散的視点は見る、聞く、学ぶなどの行動に関するものと、美術、音楽、スポーツなどの対象に関するものがある。個々の対象は各々の分野に応じて、さらに細分化した分類が可能である。位相的視点には距離あるいは地図情報に関するもの、開催時期、あるいは歴史などの時間に関するものがある。

ユーザによる検索で、すべての視点に関して項目を明確に指定する場合は目的指向の検索と呼ぶことにする。また部分的な指定に留まる場合は、ユーザの欲求が潜在化している場合もあり、制約指向の検索と呼ぶことにする。目的指向、並びに制約指向のいずれの場合も、検索するユーザ毎に様々な項目の組み合わせが生じるので、その各々に対して有意な検索条件であるかを検査する必要がある。また制約指向の場合は不完全な指定に対する対応を予め規定する必要があることが問題となる。

(2) 検索キーの整合性検査

いま視点1、2、…kに対して、各々 n_1 、 n_2 、… n_k の項目があるとする。各視点で任意に項目を指定すると、全体では $n_1 \times n_2 \times \dots \times n_k$ の組み合わせを生じる。その中には無意味な組み合わせ（例えば‘食べる’と‘美術館’）がある。このような無意味な組み合わせは、ユーザが視点間にOR条件を意識している場合に生じる。このため事前にクライアント側で検索キーの整合性検査を行って、ユーザの意図に意味のない組み合わせがある場合に、意味のある項目の組み合わせに分解することが望ましい。但し、項目数が多い場合にはすべての無意味な組み合わせを表などに保持しておくことは、メモリ量、ならびに検索の手間から事実上困難である。

この対策としては視点を2項関係に分解することが考えられる。例えば視点(1、2)、(2、3)、…(k-1、k)の関係に分解すれば、組み合わせの総数は $n_1 \times n_2 + n_2 \times n_3 + \dots + n_{k-1} \times n_k$ に削減される。ここで、2項関係として

- ＋：強い意味関係あり（例；‘食べる’＋‘食堂’）
- －：意味関係なし（例；‘食べる’－‘美術館’）
- ？：弱い意味関係あり（例；‘遊ぶ’？‘美術館’）

を定義してグラフを作成する。これは視点1、2を例にとれば n_1 、 n_2 のノードとこれらをつ結び、＋、－、？の値を持つアークからなる。

いまユーザが視点 v_1 、 v_2 、… v_j の中から項目を選択したとする。このとき選択

表 1 適合性検査のためのルール

原グラフ	等価グラフ	備考
		連続する+関係のパスは?関係と考える。
		パスの中間に-関係があるときは、全体を-関係と考える
		パスの中間に?関係があるときは、全体を?関係と考える
		単一のアーク (例では BC) と並行するパスは無視する。
		分岐、合流する並行パスでは-関係を含むパスを無視する。
		複数のアークが両端のノードを共有するとき +と? では+ -と? では- +と- では× (矛盾)と考える。
		複雑なグラフも最短経路のパスに分解して、上記のいずれかのルール (この場合は並行する同一経路長のパス) を適用する。

(注) 視点 A、C 間に選択されない視点 B があるとき、AB、BC 間に各々+関係があると考え
る。また 2 項関係は視点 (1, 2) (2, 3) … (k-1, k) のほかに (1, 5) のよ
うに視点間をバイパスするものも存在すると仮定した。() 内の数字はパス長。

された項目の整合性 (意味のある組み合わせ) は次のように検査する。

[整合性検査]

1. 選択された項目をノード、項目間の+-?関係をアークとするグラフを作成する。
2. グラフに表 1 のルールを適用して、最終的に 2 個のノードを単一アークで結んだグラフにまで簡約化する。
3. 得られた単一アークのグラフが+または?関係のとき、選択された項目は整合している。また-関係のときは整合していない。 □

ここでいくつかの注意点がある。その一つは連続する 2 つの+関係が等価的に?関係になることである。例えば '見る' と '美術' 並びに '美術' と '学校' は各々、強い意味関係を持つが、'見る' と '学校' は意味関係があいまいである。このように推移則は成り立たない。また、並行する 2 つ以上のパスがあるときは、+関係で結ばれたパスを優先して考える。但し、単一アークで結ばれている場合は、そのアークを優先する。

整合性検査の結果、等価グラフが-関係となる場合、ユーザの検索した項目を複数のグループに分割して、+または?関係の等価グラフが得られるようにする。ここで-関係は等価グラフを作成する過程で、直列のパスに-関係のアークを含むか、並列パスの一つが

-関係のアーキとなっている場合である。このことに着目して、グラフの分割は次のように行う。

[グラフの分割]

1. 等価グラフを作成する過程で①グラフがカットセットで分割され、そのカットセットが単一のアーキであるか、または両端のノードを共有する複数のアーキである、②カットセットのなかに-関係のアーキを含む、の2条件を満たすグラフを見つける。
2. カットセットに含まれる-関係のアーキと対応する原グラフのアーキ e (当然、-関係である)を見つける。アーキ e の両端のノードが x 、 y であるとする。
3. 原グラフを、ノード x を含み+関係で連結される部分グラフ X と、ノード y を含み+関係で連結される部分グラフ Y に分割する。 □

(3) 検索並びに結果の表示

観光案内における離散的視点の主たる役割は、興味の対象とする施設（動物園など）やイベント（祭りなど）を明らかにすることである。目的指向のユーザには、施設やイベントを予め視点として定義して、直ちに項目指定させることも可能である。一方、位相的視点の主たる役割は、特定の施設やイベントを限定することである。

離散的視点に対応して、サーバのデータ構造は図3のようにになっている。検索画面では複数の視点を任意に選択できたが、データベース内では図示したように固定化された階層構造をなしている。ここで各階層内の項目名はオブジェクト指向データベースにおけるクラスに対応している。このとき2項間の+関係はクラス間にリンクが存在することを意味する。また階層の根から葉に向かう+属性のリンク列はコンテンツに向かうパスに相当する。逆に、2項間の-関係はクラス間にリンクが存在しないことを意味する。また2項間の?関係は検索条件の整合性判定にのみ必要な概念で、データ構造とは直接、結び付かない。

データベースには個々の施設（××美術館など）の情報をコンテンツとして登録している。コンテンツは施設・イベントのクラスのインスタンスである。サーバの負荷を軽減するためには、少なくとも施設・イベントに至るパスをクライアント側で生成することが望ましい。これは上述したように+関係のリンクを辿ることで達成できる。例えば図3で‘学ぶ’と‘美術’が指定されたとき、‘美術’から視点（施設・イベント）に向かうパス、並びに‘学ぶ’から視点（施設・イベント）に向かうパスを辿る。このとき、両パスは‘美術学校’を共有する。従って‘学ぶ’-‘美術学校’がデータベースアクセスのための検索キーである。

位相的視点は距離や時間などの連続量を指定するためのものである。これに関しては未検討であるが、住所の一部を指定してコンテンツを限定するのであれば、離散的視点と同様な実現方法となる。図4に試作システムでの検索結果画面を示す。

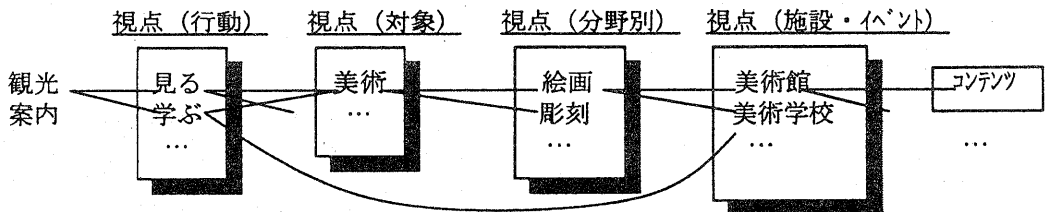


図3 離散的視点に対応するデータ構造

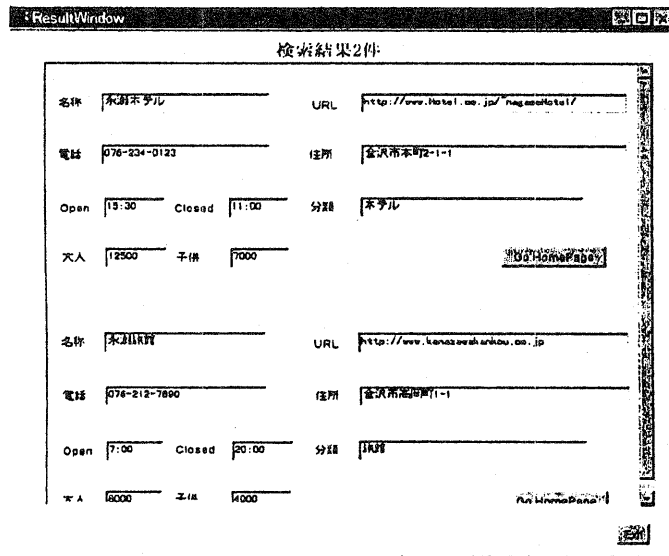


図4 検索結果の画面例

3. 実験システムの設計

ここまで述べてきた方針に従い、部分試作を行った。本システムは設備上の関係からクライアントとサーバを1台の SUN ワークステーション上に実装した。サーバには Java 言語との親和性がよいこと、フリーのオブジェクト指向データベースであることから ObjectStore^{TM(5)}を用いた。全体はシステムの流通性を考慮して、Java 言語で記述しており、クライアントとサーバ間の通信には Horb⁽⁶⁾を使用した。

試作システムはコンテンツの登録、ならびに検索機能を有する。登録は図5に示すようなテンプレートの空欄を埋めることにより実施する。施設・イベントの視点で該当する項目を選択することで、所要のテンプレートが提示される。視点間の+?関係は予めシステムに登録されているので、入力不要である。

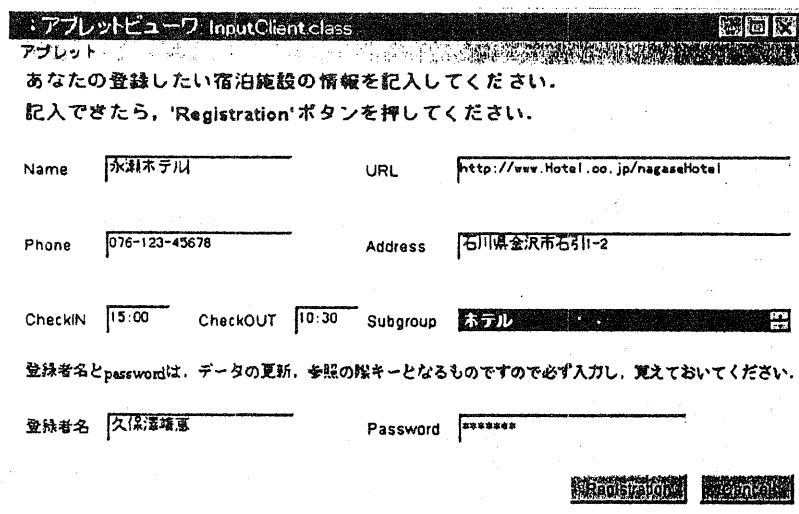


図5 登録画面例

検索も同様に図2のようなテンプレートに記入することで実施する。入力項目の整合性検査と検索キー作成はクライアントにダウンロードしたJavaプログラムにより実行する。但し、現時点では+関係しか入力できず、従って項目間が+関係で連結していれば整合すると判断している。クライアント側に離散的視点に関する処理を分担させたため、サーバ側に残る処理は①地図情報など規模の大きい処理が必要な位相的視点に関する処理、②検索キーに基づくコンテンツの検索、③データベースが分散した場合の連携処理になると考えられる。

試作システムについてはJavaのソースコード行を計数した。その結果、主としてクライアント側での動作に必要な部分が約2300行（各視点のクラス定義に約100行、コンテンツ登録・読み出しに約120行、ウィンドウ構成に約1800行、検索条件の整合性検査に約100行、データベース初期化に約180行）、サーバでの動作に必要な部分が約450行（データベース管理）であった。ユーザが検索を実行したときのHorb通信回数は最小の2回（検索条件の提示、検索結果の返送）である。開発に要した工数は約4人・月であった。この点については別稿⁷⁾で報告する予定である。本稿に述べた機能をCGIで実現する場合は①選択された視点の組み合わせに応じた動的な入力画面の編集がサーバで必要になる、②検索条件の整合性検査やグループ化をサーバで行う必要がある、③制約指向の検索では検索条件が最終的に確定するまでにクライアント・サーバ間の通信回数が増大する、等の問題点が生じる。

4. あとがき

地域内の観光情報を提供する目的でWWWのコンテンツ登録、並びに検索を支援するシステムを提案した。目的を限定した分、登録される情報の枠組みが予め規定でき、視点からの検索というユーザへの負担が少ないインタフェースを提供できた。また、システムの実装にJava言語を用いた結果、クライアントにプログラムをダウンロードして実行するという特性を生かして、サーバの負荷を軽減できた。このことは検索に伴う応答性能の向上にも寄与している。

本稿に述べてない主要な検討課題は①制約指向の検索、②位相的視点からの検索、③分散データベースへのコンテンツ格納、の3点である。例えば、今回の方法では検索条件に係る処理がデータベースの内容を色濃く反映したものになっている。上記、③分散化ではサーバに密着したアルゴリズムをどのようにクライアントにダウンロードするか、というような問題が直ちに生じる。これらについては今後、引き続き検討を進めたい。

謝辞 本研究の一部は(株)大興電機製作所の委託研究費により実施した。ここに関係各位に謝意を表す。

* ObjectStore は米国 ObjectDesign 社の商標である。

参考文献

- (1) 嶋、永瀬 “観光案内のためのコンテンツサーバ構成法の検討” マルチメディア通信と分散処理 79-4, pp.19-24, 1996
- (2) <http://www.yahoo.com>
- (3) <http://www.nsknet.or.jp/ishikawa>
- (4) <http://www.mapion.co.jp>
- (5) 石塚圭樹 “オブジェクト指向データベース” アスキー、1996
<http://www.odi.com>
- (6) <http://www.njk.co.jp/otg/Study/horbintro>
- (7) 永瀬、久保澤、渡辺 “学生による Web コンテンツ収集実験システムの試作と評価” 第2回 HORB シンポジウム 1998, 5