

地域情報サービスのための拡張 Web 空間

平松 薫[†] 石田 亨^{††}

本論文では、地域コミュニティに向けた情報サービスを実現するため、WWWを地理情報システムによって拡張した拡張 Web 空間とその検索言語を提案する。拡張 Web 空間は、ホームページ間に定義されているハイパーリンクによって構成される Web 空間を、利用者の検索文に出現するホームページ間の地理的関連性によって、動的に拡張可能としたものである。この地理的関連性を地理的ジェネリックリンクと呼ぶ。拡張 Web 空間の情報検索では、検索空間を半構造データとしてとらえ、各々のホームページに対する条件と、ホームページ間のハイパーリンクおよび地理的ジェネリックリンクに対する条件を、繰返し表現を可能にした検索パス式によって系列化して問合せを表現する。我々は、拡張 Web 空間とその検索言語をデジタルシティ京都に適用し、WWWに基づいた地域情報サービスを構築している。デジタルシティ京都では、3D ウォークスルーおよび地図インタフェースから、拡張 Web 空間に基づいた情報検索が可能となっており、本論文ではその情報検索システムとインタフェースの構成についても述べる。

An Augmented Web Space for Regional Information Services

KAORU HIRAMATSU[†] and TORU ISHIDA^{††}

We propose an augmented Web space and its query language to support geographical querying and creating sequential plans utilizing a digital city that is a city-based information space on the Internet. The augmented Web space involves a new approach to integrate the World Wide Web and a geographical information system. The augmented Web space consists of home pages, hyperlinks, and geographical generic links that represent geographical relations between home pages. The generic links are created dynamically using geographical evaluation functions included in a user's search query each time one is issued. A query also includes a path expression showing how to navigate the HPs, hyperlinks, and geographical generic links. Since the path expression is an extended regular expression, we can describe an arbitrary sequence of users' search actions for navigating the augmented Web space. We have applied the proposed augmented Web space to Digital City Kyoto, a city information service system which is accessed through a 3D walk-through and a map-based interface. Each time a user's query is issued through the 3D and 2D interfaces, Digital City Kyoto creates an augmented Web space, and navigates the Web information space based on the path expression in the query.

1. まえがき

携帯電話が高度化し、都市生活におけるインターネット利用が容易になりつつある。携帯電話が情報機器化し、対話機能を持つインタフェースエージェントが搭載される日も遠くないだろう。一方で、地理情報システム (GIS) や World Wide Web (WWW) などを用いた都市情報の集積が進んでいる¹⁾。これらを統合すれば、都市生活を支援する高度な情報サービスを構築できる。以下のシナリオを考えてみよう。

「今日は土曜日で、太郎は花子と夕食の約束をしている。しかし、その前に太郎は大学で資料を完成させ、郵便局から発送したいと思っている。太郎は携帯電話の中のエージェントに問い合わせる。『大学の近くにあつて、土曜日でも営業している郵便局を探してくれないか?』。エージェントが探している間に、太郎は花子と連絡をとる。花子は待ちきれなくて近くまで来ると言っている。太郎はエージェントに話しかける。『レストラン情報サイトで紹介されているようなイタリア料理のおすすめ店のなかで、ワインがしっかりしていて、今探している郵便局から 1 km 以内にある店を探してくれないか?』」

WWWには数多くのホームページ (HP) が存在し、HP 内部に記述されたハイパーリンクによって複雑に

[†] NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories
^{††} 京都大学大学院社会情報学専攻
Department of Social Informatics, Kyoto University

連結しあっている。また GIS には、都市のオブジェクトの位置関係が格納されている。WWW と GIS は独立したものであり、各々を対象とした検索技術は数多く提案されている。それら既存の検索技術を組み合わせることができれば、WWW の HP をそれ自身の論理的なハイパーリンクだけでなく、GIS を利用して実際の都市に即した HP 間の位置関係で関連をたどるような情報検索が可能となり、上記のシナリオの問合せが実現できると考えられる。

たとえば、『大学の近くにあつて土曜日も営業している郵便局』を探すには、まず大学のサイトから住所を抽出する。その大学の住所をキーにして、その近くにある郵便局を範囲を設定して GIS で検索する。さらに、GIS で検索した郵便局を WWW で検索し、その HP の中から営業時間に関する情報を抽出し、土曜に営業しているかどうかを判断する。『レストラン情報サイトで紹介されているイタリア料理のおすすめ店で、ワインがしっかりしていて、郵便局から 1 km 以内にある店』を探すには、レストラン情報サイト内のイタリア料理のおすすめ店のリストから各店舗の HP を検索し、ワインに関して言及のある店舗を選択する。そして選択した店舗の HP から住所を抽出し、その住所をキーにして GIS を検索し、郵便局からの距離が 1 km 以内のレストランを求めることができる。

そこで、上記のような実際の都市に即した情報検索を実現するために、WWW を GIS によって拡張した拡張 Web 空間を提案する^{*}。拡張 Web 空間は、HP 間に定義されているハイパーリンクによって構成される本来の Web 空間を、利用者の検索文に出現する HP 間の地理的関連性によって、動的に拡張可能としたものである。この地理的関連性は、HP の持つ地理的属性を GIS で評価して生成する HP 間のリンクであり、これを地理的ジェネリックリンクと呼ぶ。また、拡張 Web 空間の検索では、検索空間を半構造データ^{2),3)}としてとらえ、各々の HP に対する条件と、HP 間のハイパーリンクおよび地理的ジェネリックリンクに対する条件を、繰返し表現を可能にした検索パス式によって系列化して問合せを表現する。

本論文では、まず、WWW、GIS およびインターネット上の地図アプリケーションを対象とした従来の研究について述べる。次に、拡張 Web 空間のアイデアを提案し、そのキーである地理的ジェネリックリンクと検索パス式について述べ、拡張 Web 空間における

検索を表現するための検索言語の定義と、その検索言語を利用した検索の記述例を示す。そして、拡張 Web 空間を利用して構築した情報検索システムの構成と、デジタルシティ京都における運用、および Web ブラウザ上で動作するインタフェースの例を示し、最後に今後のシステム開発の方向性について考察する。

2. 従来の研究

2.1 半構造データベース

WWW を対象とした検索の言語表現は、比較的新しい研究分野である。データベースの分野におけるその代表的なものに、WWW を広域に展開したハイパーテキストとしてとらえたモデルを基本にした W3QL^{4),5)}と、ノードやリンクに属性を追加したグラフ構造とリレーショナルモデルを利用した WebSQL^{6),7)}がある。WWW は、タグとリンクにより構造化されていると考えることもできるが、データモデルとしては構造が不規則なデータの集まりであり、はっきりしたスキーマが定義できないため、半構造データとして考えた方が適当と考えられている^{2),3)}。

W3QL、WebSQL は、ともにこうした半構造データモデルを前提とし、SQL を拡張した検索言語を提供している。構造の分からない Web 空間から必要な情報を検索するためには、HP に記述された情報に対する条件や HP 間のハイパーリンクに対する条件の評価だけでなく、検索経路の指定が重要となる。特に構造の分からない Web 空間から住所などの地理的属性を探し出すためには、検索経路の指定に正則表現に基づく繰返し表現を利用することが有効と考えている。

2.2 地理的関係演算

地理的関連性を求める関係演算は、Open GIS Consortium⁸⁾によって標準化が行われており、2次元もしくは3次元の座標系を利用したオブジェクトモデルやオブジェクト間の関係演算、データベース操作のための SQL の拡張が提案されている。都市のオブジェクトは、位置、形状、住所、所有者といったオブジェクトの属性に応じて構造化することが可能である。また、地理的関係演算を利用すると、オブジェクト間の距離や相対位置関係、包含、接続、交差などを評価することができる。

従来の WWW 検索では、HP 間に潜在する地理的関係の評価して情報検索に反映させることは困難であったが、情報の持つ地理的属性を評価できるようにすることで、都市に即した情報検索が可能になると考えている。

^{*}ただし本論文では、エージェント技術、携帯端末技術に関する詳細な議論はしない。稿を改めて報告する。

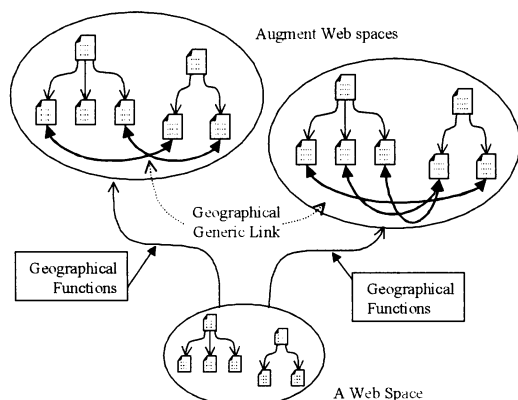


図 1 拡張 Web 空間の位置付け

Fig. 1 Conception of augmented Web spaces.

2.3 地理情報システムと WWW の結合

ネットワークを利用して情報を流通させる地理情報システム (GIS) が普及するのにもない、インターネット経由で利用できる地図情報サービスが数多く提案されてきた。その代表例としては、マビオン⁹⁾、マップファンウェブ¹⁰⁾、JaMaPS¹¹⁾などがある。これらのサービスでは、地図上に表示されたアイコンから関連情報の閲覧が可能となっているだけであるが、今後は、地図情報と WWW を融合した情報検索が、インターネットにおいて重要な役割を持つようになると考えている。

情報検索の側面から考えた場合、地図アプリケーションは検索結果の一覧性において優れているものの、地図上で可能な操作は表示された情報の直接選択が中心であるため、WWW のハイパーリンクをたどって情報を探するような系列的な情報検索の実施は困難となっている。地理情報と WWW を組み合わせてより複雑な情報検索を可能にするためには、インタフェースの改善に加え、検索システム内部の動作を規定するためのモデルの構築が重要になると考えている。

3. 拡張 Web 空間

3.1 基本的なアイデア

本論文で提案する拡張 Web 空間は、半構造データである Web 空間を、地理的ジェネリックリンクで動的に拡張したものである (図 1)。実際の都市で HP 間に地理的な関係があったとしても、ハイパーリンクで構成された Web 空間にはその関係が反映されているとは限らない。そこで、HP 間の地理的な関係を表す地理的ジェネリックリンクを動的に生成することによって、実際の都市に即した形に Web 空間を拡張する。

拡張 Web 空間における情報検索は、W3QL^{4),5)}お

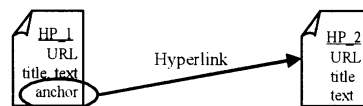


図 2 ハイパーリンク

Fig. 2 Hyperlink.

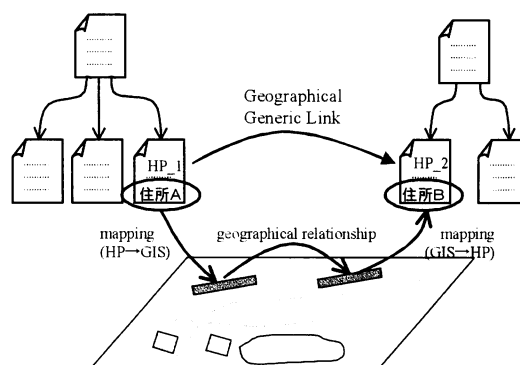


図 3 地理的関連性を用いた地理的ジェネリックリンクの例

Fig. 3 An example of a geographical generic link by geographical relationship.

よび WebSQL^{6),7)}に代表される WWW を半構造データとして扱う検索言語に、実際の都市におけるオブジェクト間の地理的関連性を評価するための地理的関係演算を導入して実現する。

HP に対する条件と、HP 間のハイパーリンクおよび地理的ジェネリックリンクに対する条件は、繰返し表現を含んだ検索ハス式で系列化する。繰返し表現を導入することで、半構造データである拡張 Web 空間に適応した検索経路の表現が可能となる。また、検索経路上の条件は、幅優先探索などの探索アルゴリズムに基づいた順序で評価を実施する。

情報空間の拡張としては、GIS を WWW のハイパーリンクで拡張する拡張 Web 空間とは反対のアプローチも可能ではあるが、拡張した情報空間を実際の都市にあわせて収束させることと、インターネット上に構築するデジタルシティ京都からの利用を考慮し、本論文では WWW を GIS で拡張する拡張 Web 空間のアプローチを採用する。

3.2 地理的ジェネリックリンク

拡張 Web 空間では HP 間のリンクとして、ハイパーリンクに加えて、地理的ジェネリックリンクを定義する。

ハイパーリンク (図 2) が、HP 内に記述された URL による HP 間の有向リンクであるのに対し、地理的ジェネリックリンク (図 3) は、HP 間に潜在する地理的関連性を検索時に GIS 内で評価して生成する動的な

リンクを指す。たとえば、「デパート A の西隣にあるデパート B」、「市役所から 100 m 以内にあるバス停」といった都市におけるオブジェクト間の地理的関連性を利用して、HP 間に地理的ジェネリックリンクを生成することができる。

地理的ジェネリックリンクは、図 3 に示すように、HP から GIS 内の地理的なオブジェクトへのマッピング、GIS を利用したオブジェクト間の地理的関係演算の実施、GIS 内のオブジェクトから HP へのマッピングの 3 ステップを経て生成される。HP から地理的なオブジェクトへは、HP から住所や店舗名などの地理的属性を抽出し、その属性に対応するオブジェクトを GIS で検索してマッピングを行う。次に、HP からマッピングされた地理的なオブジェクトを起点に、距離などを条件とした地理的関係演算を実施し、指定された条件に合致するオブジェクトを検索する。そして、検索した GIS 内のオブジェクトの住所、ビル名、所有者といった属性に対応する HP を WWW で再度検索して、地理的ジェネリックリンクを生成する。

拡張 Web 空間では、地理的ジェネリックリンクを地理的関係演算に基づいて生成可能なものとして考えるが、実際の WWW および GIS を対象にした場合には、記載情報の不完全性や語彙の不一致などより、HP から GIS 内のオブジェクト、および GIS 内のオブジェクトから HP へのマッピングで、オブジェクトの属性すべてが一致するとは限らない。この問題に対応するため、本論文の地理的ジェネリックリンクの生成では、抽出した属性間の一致条件を順次緩めながら繰り返し検索を行い、オブジェクトのマッピングを実施する。

明示的に記述されたリンク以外を導入した先行研究としては、ハイパーテキストデータベースシステムを対象として、ドキュメントに存在するリンクとは独立にクエリの組合せによるリンク (query-pair link) に変数を導入した TextLink-III¹²⁾、TextLink/Gem¹³⁾がある。query-pair link は、2つの任意のクエリを組み合わせてリンクを生成するのに対し、地理的ジェネリックリンクは、HP に対応する GIS 内のオブジェクト間の地理的関係によってリンクを生成する。また、地理的ジェネリックリンクの場合は、WWW と GIS の間で属性の表現が異なる場合にもオブジェクトのマッピングができるように、属性間の一致条件の制御を実施しながらマッピングを実施する。

3.3 検索パス式

拡張 Web 空間では、HP 間のリンクに関する条件の系列を繰り返し表現を含んだ検索パス式で指定し、検

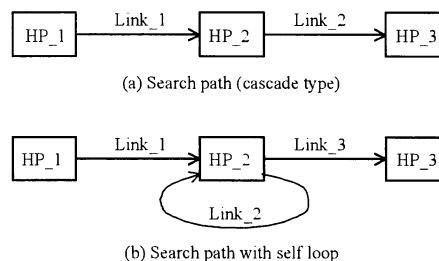


図 4 検索パス式による検索パスのパターン

Fig. 4 Path patterns by search path expression.

索を行う。拡張 Web 空間における検索パス式は、半構造データベースを対象とした検索言語 W3QL のパターングラフ (pattern graph) を基本にしたものである。

検索パス式では、起点となる HP から終点となる目的の HP までの検索経路を、HP とリンクの繰返しによって記述する。たとえば、図 2、図 3 の検索パス式はそれぞれ、“HP_1, HyperLink, HP_2”、“HP_1, GeographicalGenericLink, HP_2” と記述する。検索パス式の中では、ハイパーリンク、地理的ジェネリックリンクともに同様に記述できるものとする。

3 つ以上の HP を経由する検索経路は、図 4 の形を基本に構成する。図 4 (a) は図 3 の例にさらにリンクと HP を接続したものであり、その検索パス式は“HP_1, Link_1, HP_2, Link_2, HP_3” と記述する。

これに対し図 4 (b) は、Web 空間の検索で目的の情報の正確な位置が予測できない場合に対応するため、繰返し表現によってパス中の HP にそれ自身へのセルフループを設定した例である。検索式は“HP_1, Link_1, (HP_2, Link_2), Link_3, HP_3” と記述し、括弧でくくった“(HP_2, Link_2)”の部分で繰返しを表現する。この表現を利用すると、たとえば、内部構造の分からないレストラン情報サイトのトップページから有限回リンクをたどってワインに関する言及のあるイタリア料理店を探すといった検索表現を、起点の HP から有限回リンクを経由して目的の HP まで到達するという形の検索パス式で記述することが可能となる。

4. 検索言語

本章では、3 章で述べた拡張 Web 空間における検索を表現するための検索言語の文法を定義し、その意味論と記述例を示す。

4.1 文法

本論文で扱う拡張 Web 空間の検索言語は、W3QL^{4),5)} を基本に、HP の属性の扱いと評価方法

```

query ::= SELECT select_clause
        FROM from_clause
        WHERE condition+
select_clause ::= object+
object ::= homepage | homepage.attribute
attribute ::= url | server | title | anchor | href | text | address
from_clause ::= path *
path ::= step * homepage
step ::= homepage, link, | (homepage, link), link,
condition ::= homepage.attribute+ operator value | link+ AS link_type
value ::= True | False | numerical_value |
         string_value | regular_expression | homepage.attribute
link_type ::= HYPER_LINK | geographical_generic_link
geographical_generic_link ::= gis_function( homepage, homepage ) operator value
gis_function ::= Equals | Disjoint | Touches | Within | Overlaps | Crosses |
               Intersects | Contains | Distance

```

図5 検索言語の文法

Fig.5 Syntax of query language.

を変更し、HPの持つ地理的属性を検索式中で扱えるように地理的ジェネリックリンクを導入したものである。検索式は、SELECT-FROM-WHERE形式で記述するものとし、図5のように定義する。

定義中の大文字列で始まるトークンは終端記号、小文字列は非終端記号、`*`は0回以上の繰返し、`+`は1回以上の繰返しを意味する。`value`は、数値、文字列、文字列の正規表現、もしくは検索パス中に出現するHPの属性を表現する。`operator`には、Perl言語で定義されている2項演算子を利用するものとし、HPの属性と`value`の部分一致、完全一致、もしくは大小関係を評価するとき用いる。

4.2 意味論

`select_clause`には、検索結果として出力する項目をHP(`homepage`)もしくはHPの持つ属性(`homepage.attribute`)で指定する。`from_clause`には、検索の起点のHPから終点のHPまでの経路を、3.3節で述べた検索パス式(`path *`)で記述する。`homepage`と`link`の繰返しは括弧を用いて表現し、Web空間の半構構性に対応した検索経路を実現する。

`where_clause`には、検索パス式上に出現するHPもしくはその属性に対する条件、およびHP間のリンクの種類とリンクに対する条件を記述する。記述された条件は、`from_clause`に記述した検索パス式の経路に従って評価する。同じHPもしくは同じリンクに対して複数の条件が列挙されている場合は、それらを条件の積集合として処理する。

表1 地理的ジェネリックリンクで用いる地理的關係演算例

Table 1 Examples of GIS functions for creating geographical generic links.

関数名	意味	戻り値
Equals	同値	True/False
Disjoint	共通部なし	True/False
Touches	接触	True/False
Within	内部に存在	True/False
Overlaps	重複部あり	True/False
Crosses	交差	True/False
Intersects	共通部あり	True/False
Contains	包含関係	True/False
Distance	距離	連続値

リンクに関する条件は、まずASによってリンクがハイパーリンクであるか、地理的関連性によって生成する地理的ジェネリックリンクであるかを指定する。ハイパーリンクの場合は、リンク前者のHPの内部に記述されている情報に従ってリンク後者のHPが展開される。また、地理的ジェネリックリンクの場合には、リンク前者のHPに対応する地理的属性を起点に、指定された地理的関連性を持つオブジェクトをGISで検索し、検索したオブジェクトに対応するHPがリンク後者のHPとして展開される。地理的関連性に関する条件は、地理的関係演算の種類により、数値で評価する場合と、真偽で評価する場合に分けられる(表1)。

4.3 実行例

(1) 地理的ジェネリックリンクを用いた検索

起点となるHPに対応する建物から100m以内に存在する建物のHPを、地理的ジェネリックリンクを用

いて検索する例を示す。

```
1 SELECT target.url
2 FROM start,distance_link,target
3 WHERE start.url eq 'http://www.xx.yy/'
4 distance_link AS
    Distance(start,target) < 100m
```

1 行目で目的の HP の URL (target.url) を検索結果として出力するように記述, 2 行目で起点の HP (start), 地理的ジェネリックリンク (distance_link), 目的の HP (target) からなる検索パス式を記述している。3 行目で起点となる HP の URL を指定, 4 行目で distance_link が距離条件による地理的ジェネリックリンクであることを記述し、距離が 100m 以内という条件でリンクを生成する。

この検索例では、target に対する条件が設定されていないので、地理的ジェネリックリンクだけが条件として評価され、100m 以内に存在する GIS オブジェクトに関連した HP がすべて列挙されることになる。

(2) 繰返し表現を含んだ検索式

検索パス式中に、図 4 (b) で示した繰返し表現を含んだ検索例を示す。

```
1 SELECT tel.url
2 FROM start,link,(middle,link),link,tel
3 WHERE link AS HYPER_LINK
4 start.url eq 'http://www.xx.yy/'
5 middle.host eq 'www.xx.yy'
6 tel.text =~ /telephone/
```

1 行目で、検索結果として目的の HP の URL (tel.url) を出力するように記述, 2 行目には“(middle,link),”と繰返し表現を含んだ検索パス式を指定する。検索パス式上のリンク link は HP 内に記述されたハイパーリンクであり (3 行目)、URL で指定された起点の HP (start, 4 行目) から、同じホストに存在する HP へと接続したリンクをたどり (5 行目)、“telephone”という文字列を含む目的の HP (tel) を検索する。

繰返し表現を導入することで、情報サイト www.xx.yy の内部構造が正確に分からない場合でも、“telephone”というキーワードを含む HP をすべて検索することができる。

(3) 例題

1 章のシナリオは、GIS を利用して大学の近くの郵便局の検索し、検索した郵便局に対応する HP を調べて営業時間の条件に合う郵便局を検索すること、検索した郵便局から 1 km 以内にある、レストラン情報サイトで紹介されているイタリア料理のおすすめ店のう

ち、ワインに関する言及のある店を検索することに分けることができる。

前半の「大学の近くにあつて土曜日にも営業している郵便局」の検索では、まず起点を大学に設定し、GIS で大学の近くにある郵便局を検索し、検索した郵便局の HP を WWW で検索して、その中で営業時間に関する条件が合うものを結果として出力する。この検索では、「近く」という条件を距離 500m と設定し、以下のように検索式を記述する。

```
1 SELECT schedule.url
2 FROM university,distance_link,
    (postoffice,link),link,schedule
3 WHERE link AS HYPER_LINK
4 university.url eq 'http://www.zz.edu/'
5 distance_link AS
    Distance(university,postoffice) < 500m
6 postoffice.title =~ /postoffice/
7 schedule.text =~ /open on saturday/
```

1 行目で検索結果として出力する schedule.url を指定, 2 行目でその検索パス式を指定する。検索パス式では、まず、大学の HP (university) を起点に、500m 以内という距離条件に基づく地理的ジェネリックリンク (distance_link, 5 行目) と HP のタイトルにキーワードである郵便局 “postoffice” が含まれるという条件 (6 行目) で郵便局の HP を検索する。そして検索パス式後半“(postoffice,link),link,schedule” は、郵便局サイトの内部構造にかかわらず、ハイパーリンクに従い、そのサイト全体の検索を行う。7 行目の営業時間に関するキーワード “open on saturday” の含む HP を発見した場合に、結果を schedule.url として出力する。

後半のイタリア料理店の検索では、レストラン情報サイトのイタリア料理のおすすめ店リストを起点に、先に検索した郵便局との距離が 1 km 以内で、イタリア料理店のメニューの中でワインに関して言及のある店を検索結果として出力する。

この検索は、以下のように記述する。

```
1 SELECT restaurant
2 FROM info_top,link,(info,link),link,
    restaurant,distance_link,postoffice
3 WHERE link AS HYPER_LINK
4 info_top.url eq 'http://www.info.com/'
5 info.host eq 'www.info.com'
6 restaurant.text =~ /italian/
7 restaurant.text =~ /wine/
8 postoffice eq 'http://www.post.com/'
```

9 distance_link AS

$\text{Distance}(\text{restaurant}, \text{postoffice}) < 1\text{km}$
 1 行目で、検索結果として `restaurant` を指定し、2 行目でその検索パスを指定する。レストラン情報サイトのトップページ `info_top` を 4 行目で指定し、検索式括弧内の繰返し表現と 5 行目の条件で、サイトの内部構造にかかわらず検索パスを生成し、キーワード“italian”を含むイタリア料理店でワイン“wine”に関する記述のある HP (`restaurant`) を検索する。そして、前に検索した郵便局サイトとの間で距離を条件とした地理的ジェネリックリンクを生成し、条件に合致するレストランサイトの URL を検索結果として出力する。

5. システム構成

拡張 Web 空間に基づいた情報検索システムは、図 6 に示すように、検索制御モジュール、Web モジュール、形態素解析器、GIS、地理的ジェネリックリンク生成モジュール、そして地理的ジェネリックリンクのキャッシュモジュールとそのデータベースで構成する。なお、本情報検索システムはワークステーション[☆]上で開発を行い、形態素解析器には茶筌^{☆☆}を利用、GIS には PostgreSQL と組み合わせた住宅地図データベース^{☆☆☆}を利用している。また、PostgreSQL は地理的ジェネリックリンクをキャッシュするデータベースとしても利用している。それ以外のモジュールは Perl で実装を行っている。

検索制御モジュールは、インタフェースからのクエリを受け付け、検索を実行する。まず、*where_clause* に記述された HP に対する条件と HP 間のリンクに関する条件を抜き出し、*from_clause* に記述された検索パス式上の各要素に条件をバインドする。そして検索制御モジュールは、Web モジュールおよび地理的ジェネリックリンク生成モジュールと連携し、検索パス式に指定された検索経路上の HP の条件評価とリンクの生成・展開を実施し、クエリに対応する目標の検索を行う。検索結果は、*select_clause* に記述された出力形式に従って出力される。

Web モジュールは、検索制御モジュールからの指示に従い、HP の属性に対する条件評価と、ハイパーリンクの抽出を行う。このうち HP の属性およびハイパーリンクの抽出は、実際にインターネットにアクセスして指定された HP を取り込み、HTML タグに基

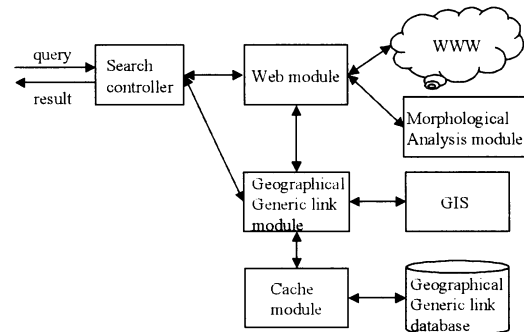


図 6 システム構成

Fig. 6 System architecture.

づいた HP の解析とテキスト部の形態素解析を行って実施する。

5.1 地理的ジェネリックリンクの生成とキャッシュ

地理的ジェネリックリンク生成モジュールは、検索制御モジュールからの指示に従い、HP 間の地理的な関係を表す地理的ジェネリックリンクを生成する。地理的ジェネリックリンクは、3.2 節で述べたように、HP から GIS 内の地理的なオブジェクトへのマッピング、GIS を利用したオブジェクト間の地理的関係演算の実施、地理的なオブジェクトから HP へのマッピングの 3 ステップを経て生成される。なお、現在の情報検索システムの実装では、地理的な座標系において点で表現可能なオブジェクトを利用して、HP と GIS 内のオブジェクトの間の相互マッピングを行っている。

(1) **HP から地理的なオブジェクトへのマッピング**
 Web モジュールを利用してリンクの起点となる HP から住所や店舗名などの地理的属性を抽出し、起点 HP に対応する地理的なオブジェクトを GIS で検索する。この際、HP から抽出した地理的属性が GIS のオブジェクトの属性と完全に一致しない場合があるため、抽出した属性すべての完全一致から住所や店舗名の部分一致まで、検索されたオブジェクト数を基準に条件を緩めながら、繰り返し検索を実行する。

(2) 地理的関係演算の実施

起点 HP に対応する GIS 内のオブジェクトから、検索制御モジュールから指定された地理的関連性にあるオブジェクトを GIS で検索する。運用中のシステムでは、点で表現可能なオブジェクトを利用してマッピングを行うことを考慮し、表 1 の関数のうち Within と Distance を実装している。

(3) **地理的なオブジェクトから HP へのマッピング**
 検索した GIS 内のオブジェクトの属性をキーワードとして、対応する HP を検索する。この際、GIS 内のオブジェクトに対応する HP が存在しない場合がある

[☆] Sun Enterprise250: Sun Microsystems, Inc.

^{☆☆} 茶筌: 奈良先端科学技術大学院大学自然言語処理学講座

^{☆☆☆} 住宅地図データベース Zmap-TOWNII: 株式会社ゼンリン

ため、HP から GIS へのマッピングと同様に、属性すべての完全一致から住所や店舗名の部分一致まで、検索された HP 数を基準に条件を緩めながら、検索を繰り返し実行する。

地理的ジェネリックリンクの生成には、HP へのアクセスおよび GIS の検索が必要なため、後述する地図インタフェースやウォークスルーからの問合せを実用的な速度で処理することは、現在の技術では困難と考えている。この問題を解決するために本情報検索システムでは、地理的ジェネリックリンク生成時に抽出される HP の URL と GIS オブジェクトの地理的座標の組合せをキャッシュとして蓄積し、リンク生成時に再利用する。このキャッシュを利用すると、HP と GIS のオブジェクトの間の相互マッピングを省略でき、地理的ジェネリックリンク生成の高速化を図ることができる。

5.2 検索パス式の評価方法

検索制御モジュールは、検索パス式に指定された検索経路上の HP の条件評価とリンクの展開を実施し、クエリに対応する目標の検索を行う。この条件評価とリンクの展開は、探索問題の一種として考えることが可能であり、現在の実装では、アルゴリズムに幅優先探索を利用して情報検索システムを構築している。

6. デジタルシティ京都における運用

我々は、デジタルシティ京都[☆]において情報検索システムを運用し、そのインタフェースとして 3D ウォークスルーと地図インタフェースを実験的に公開している。

情報検索システムの公開にあたっては、プライバシーの保護や著作権、著作権といった社会的問題に配慮し、問題の解決した HP だけを登録したデータベースを利用して、検索結果の出力に制限をかけている。具体的には、著作権および著作権に関して HP 作成者から承諾の得られた HP だけをデータベースに登録し、正確な住所が分かることにより問題が発生する可能性のある個人 HP などに関しては登録を制限している。登録 HP のデータベースは、地理的ジェネリックリンクのキャッシュと同様の形で管理しており、現在 2563 件^{☆☆}の京都市内に関連した HP が登録されている。

6.1 地理的ジェネリックリンクの有効性

デジタルシティ京都の登録 HP は、作成者に対して確認を行い、正確な住所情報の付与も行っている。この情報を利用して、地理的ジェネリックリンク生成モ

ジュールによる HP と地理的なオブジェクトの間の相互マッピングの実験を行った。

実験では、登録されている HP 2563 件を対象に、地理的な点に基づく相互マッピングを実施した。登録 HP を単独の HP として扱い、相互マッピングを実施したところ、HP から GIS 内のオブジェクトへのマッピングの成功率が 18.9%、GIS 内のオブジェクトから HP へのマッピングの成功率が 39.0%であった。なお、実験対象とした登録 HP のうち、住所に関する表現が含まれていた HP は 68.5%、そのうち、正確な番地レベルの住所まで記載があったのは 34.8%であった。

実際の Web サイトでは、トップページ以下、HP が木構造になっている場合が多く、地理的ジェネリックリンクの作成に必要な情報を効果的に抽出するには、関連のある HP をクラスタとして扱えるような検索パス式の記述が必要と考えている。また、京都特有の住所の通称表記への対応や、対象地域特有の固有表現が抽出できるような形態素解析器の拡張も、相互マッピングの精度の改善には重要と考えている。

6.2 公開インタフェースからの検索

我々は、情報検索システムのインタフェースとして、3D ウォークスルーと地図インタフェースをデジタルシティ京都で実験的に公開している。各インタフェースは、情報検索システムに CGI プログラム経由でアクセスし、検索結果を仮想空間中、もしくは地図上に表示する。都市のメタファをインタフェースに利用することによってユーザの利便を図り、登録 HP を対象に検索を実行することによって、社会的問題への配慮と実用的な速度での動作を可能にしている。

(1) 3D ウォークスルーからの検索

米国 Flatland 社が作成した 3DML プラグインを利用して、京都市四条通の約 1.6 km を 3D ウォークスルー(図 7)で再現した。3DML は、HTML とほぼ同様の形式を持った簡易型 3 次元空間記述言語であり、格子上のマップにオブジェクトを配置していくことで、仮想的な 3 次元空間を作成することができる。この 3D ウォークスルーは、Web ブラウザ上で利用することが可能であり、ユーザは仮想空間中をマウスやカーソルキーを利用して自由に移動することができる。

また 3D ウォークスルーの路面には、下記のようなクエリを情報検索システムで実行する CGI プログラムへのリンクが埋め込んであり、ユーザの移動に合わせて、現在位置周辺の HP へのリンクがブラウザ下部に表示されるようになっている。このクエリでは、ウォークスルー内の現在位置から距離 20 m 以内の HP に対して地理的ジェネリックリンクを生成し、リンク

[☆] <http://www.digitalcity.gr.jp/>

^{☆☆} 2000 年 5 月 20 日現在



図 7 3D ウォークスルー
Fig. 7 3D walk-through.

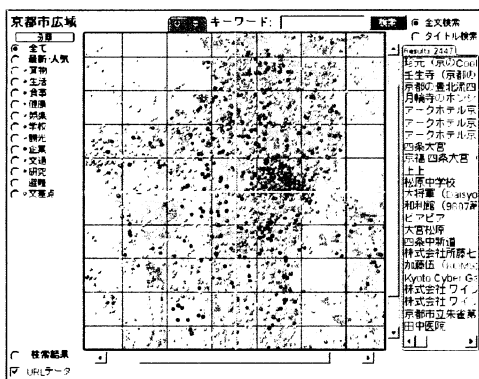


図 8 地図インタフェース
Fig. 8 Map-based interface.

先 HP の URL を検索結果として出力する。

```
1 SELECT shop.url
2 FROM position,distance_link,shop
3 WHERE position.url eq '(given by CGI)'
4 distance_link AS
```

Distance(position,shop) < 20m

(2) 地図インタフェースからの検索

地図インタフェース(図 8)からは、情報検索システムを利用して、表示している地図を検索範囲としたカテゴリ検索、キーワード検索および距離に基づいた検索が可能となっている。検索結果は、アイコンとして地図上に重ね合わせて表示され、アイコンをクリックするとその HP が別ウインドウに表示される。また、出力された検索結果を起点に、地理的属性や HP の属性に関連した情報の再検索もすることができる。

6.3 デジタルシティ京都における今後の課題

6.3.1 検索インタフェースの課題

拡張 Web 空間は、現在開発中のデジタルシティ京都で生じた実際のニーズに基づいているが、6.2 節

で述べたインタフェースからの検索はそれほど複雑にならない。これは、利用者が 3D ウォークスルーや地図インタフェースから問合せを行うため、曖昧な検索意図をインタフェース上で具体化し、単純な検索式に分解する必要があるからである。

利用者自身が検索意図を分解することなく、利用者の曖昧な検索意図を表現するような複雑な検索を行えるようにするためには、地図インタフェースの拡張や、自然言語処理が可能なインタフェースエージェントの導入が必要と考えている。

(1) 地図インタフェースの拡張

地図インタフェースは、検索結果の一覧性が高い反面、いくつか目標を設定して全体のプランを作成するような、構成的検索を行うことは困難である。たとえば、検索した建物を起点に最寄りの駅までの経路を地図インタフェースで検索する場合、まず起点の建物に印をつけて、改めてメニューから距離検索やキーワード検索を実施する必要があるなど、地図上の直感的な操作だけでは実現できない操作が存在する。このため構成的検索を実施するためには、自然言語インタフェースとの統合や、ユーザの操作順序に基づいた検索手順の提示機能などの検討が必要と考えている。

(2) インタフェースエージェントの導入

情報検索システムのフロントエンドに Exttempo のインタフェースエージェント¹⁴⁾などを利用することで、対話を通じた利用者の検索意図の抽出や、知識やそれまでの経験を組み合わせた検索式の生成が可能になると考えている。利用者が検索式を作成する場合と比較し、より多くの情報を参照しながら効果的な検索式の生成することが期待できるため、その結果、より地域コミュニティにおける活動に役立つ情報検索が可能になるだろうと予想している。

6.3.2 情報検索システム構築に関する課題

情報検索システムの現在の実装では、各インタフェースからの検索の高速化と、検索結果に関する社会的問題への配慮のために、拡張 Web 空間で実現できる検索機能を一部制限したものとなっている。今後は、検索結果に対する配慮を残しつつ、拡張 Web 空間を利用した様々な検索ができるように、検索インタフェースの拡張とあわせて、以下の課題を検討していく予定である。

(1) WWW と GIS の間のマッピングの改善

現在の実装では、点で表現可能な GIS 内のオブジェクトに基づいてマッピングを行っている。今後は、検索言語、インタフェース、および HP からの情報抽出方法を拡張することによって、道路や地域など地理的な

線や面による WWW と GIS の間のマッピングを可能にし、より広範なオブジェクトの多様な地理的關係に基づいた検索ができるようにシステムの拡張を行う予定である。

(2) 検索アルゴリズムの改善

現在の実装では、単純な探索アルゴリズムを利用して検索パス式の展開を行っている。今後は、実際の WWW や GIS に即した探索アルゴリズムの最適化や、検索パス式中の正規表現の展開方法、インタフェースの動作とあわせたクエリ作成方法の検討が必要と考えられている。

7. むすび

本論文では、WWW を GIS によって拡張した拡張 Web 空間を提案し、その検索言語の提案を行った。拡張 Web 空間では、半構造データベースの検索言語において研究された検索パス式と、標準化の行われている地理的関連演算を用いて、複雑な検索を可能とした。

実際の状況に適した情報検索システムの構築には、探索アルゴリズムの検討とともに、ユーザからの条件追加に応じてクエリを構築していくような、インタラクティブな問合せにも対応できるインタフェースの開発もあわせて検討する必要があると考えている。その他、検索式を生成するための自然言語解析機能、検索結果の集合の大きさの大きさや検索の並行実行可能性、逐次実行可能性の判断などが今後の課題と考えている。

参 考 文 献

- 1) Ishida, T., Akahani, J., Hiramatsu, K., Isbister, K., Lisowski, S., Nakanishi, H., Okamoto, M., Miyazaki, Y. and Tsutsuguchi, K.: Digital City Kyoto: Towards A Social Information Infrastructure. *Cooperative Information Agents III*, pp.23-35 (1999).
- 2) Abiteboul, S.: Querying Semi-Structured Data. *Database Theory - ICDT '97, 6th International Conference*, pp.1-18 (1997).
- 3) 田島敬史：半構造データのためのデータモデルと操作言語。情報処理学会論文誌, Vol.40, No.SIG 3(TOD 1), pp.152-170 (1999).
- 4) Konopnicki, D. and Shmueli, O.: W3QS: A Query System for the World-Wide Web. *Proc. 21st International Conference on Very Large Data Bases*, pp.54-65 (1995).
- 5) Konopnicki, D. and Shmueli, O.: Information Gathering in the World Wide Web: The W3QL Query Language and the W3QS System. *ACM*

Trans. Database Syst., Vol.23, No.4, pp.369-410 (1998).

- 6) Mendelzon, A., Mihaila, G. and Milo, T.: Querying the World Wide Web. *International Journal on Digital Libraries*, Vol.1, No.1, pp.54-67 (1997).
- 7) Mendelzon, A. and Milo, T.: Formal Models of Web Queries. *Proc. 16th ACM Symposium on Principles of Database Systems*, pp.134-143 (1997).
- 8) Open GIS Consortium, Inc: OpenGIS Simple Features Specification For SQL Revision 1.0. <http://www.opengis.org/>
- 9) マピオン. <http://www.mapion.co.jp/>
- 10) マップファンウェブ. <http://www.mapfan.com/>
- 11) 高木 悟：分散型電子地図プラットフォーム JaMaPS, 画像ラボ, Vol.9, No.12, pp.27-32, 日本工業出版 (1998).
- 12) Tanaka, K., Nishikawa, N., Hirayama, S. and Nanba, K.: Query Pairs As Hypertext Links. *Proc. 7th IEEE Data Engineering Conference*, pp.456-463 (1991).
- 13) Qian, Q., Tanizaki, M. and Tanaka, K.: Abstraction and Inheritance of HyperLinks in an Object-Oriented Hypertext Database System TextLink/Gem. *IEICE Trans. Information and Systems*, Vol.E78-D, No.11, pp.1343-1353 (1995).
- 14) Extempo Systems Inc. <http://www.extempo.com/>

(平成 12 年 3 月 20 日受付)

(平成 12 年 7 月 10 日採録)

(担当編集委員 田中 克己)



平松 薫 (正会員)

平成 6 年慶應義塾大学理工学部電気工学科卒業。平成 8 年同大学院理工学研究科計算機科学専攻修士課程修了。同年日本電信電話(株)入社。現在 NTT コミュニケーション科学基礎研究所に所属。



石田 亨 (正会員)

昭和 51 年京都大学工学部情報工学科卒業。昭和 53 年同大学院修士課程修了。同年日本電信電話公社電気通信研究所入所。現在、京都大学大学院社会情報学専攻教授。工学博士。