

不完全質問処理に基づくマルチメディアデータベース応用を記述するための言語

田渕 仁浩 村岡 洋一

早稲田大学大学院 理工学研究科

本稿では、提案中のMeSODモデルと不完全質問処理に基づくデータベース応用を記述するための言語を提案した。提案した言語の特徴は大きく分けて三つある。一つは、問い合わせ条件に予測される不完全さに対する許容の仕方をデータ定義の中で支援する点である。もう一つは、問い合わせ文に不完全質問節を導入して問い合わせ自身にも不完全さに対する言明を記述できる点である。3番目の特徴は問い合わせ条件の中でマルチメディアデータをキーとして含められるような外部関数を定めている点である。また、これらの定義は鳥類図鑑におけるマルチメディアキーの検索や大相撲放送に対する動画検索の例を用いて説明した。

A language for description of a multimedia application based on the MeSOD model and query-processing of incomplete queries

Masahiro Tabuchi, Yoichi Muraoka

Graduate school of Science and Engineering, Waseda University

3-4-1 Okubo, Shinjuku-ku, Tokyo, 169 Japan

E-mail: tabuchi@cfi.waseda.ac.jp

The purpose of this paper is to propose a language for description of a multimedia application based on the MeSOD model and query-processing for incomplete query. The language has three major characteristics. First one is to allow database administrators to describe constraints for an incompleteness of a retrieval condition into a data definition. Second one is to allow application programmers and users to describe assertion of an incompleteness of a retrieval condition in a query. Third one is to define external functions which extend a retrieval condition in a query to include multimedia keys. These definitions are illustrated in the examples of multimedia key retrieval on a database of birds and motion picture retrieval on a database of SUMO programs.

1. まえがき

本稿ではMeSODモデル[1]と不完全質問処理[2]のためのデータ定義と問い合わせ記述のための言語を提案する。

提案中のMeSODモデルではデータベースを距離空間の集まりとして定義する。MeSODモデルでは類似したデータ間の距離が近くなるような距離関数が定義されているとする。例えば、単純な数値集合や文字集合の場合には一般的な距離関数の定義で近さを表現できる。

しかし、より複雑な構造を持つデータに対して普遍的な距離関数を定義するのは困難であるために、MeSODモデルでは問い合わせ処理を工夫して、指定された条件に類似したデータを取り出す。この仕組みを不完全質問処理とよび、与えられた問い合わせ条件を書換えることによって妥当な候補を提供するような問い合わせ処理として定義している。

不完全質問処理は書換えに対する意味制約と書換え評価基準をもとに問い合わせ条件を書換える。書換えられた問い合わせ条件はデータベース中のあるデータもしくはデータ集合に対応し、それがもとの問い合わせ条件の候補になる。

書換えに対する意味制約は、MeSODモデルで定義された距離空間内の2点間の距離値として計量される書換えコストに関連づけられている。つまり、書換えコストの値に対する制限と制限を越えたときの書換え操作の継続の可否判定を意味制約として与えられる。

また、異なる空間同志の書換えコストの価値の大小を書換え評価基準として与えられることによって、良い書換えの基準を書換えコストの安さに還元できる。

これらの定義の結果、目的のデータに対する典型的な条件から派生可能な問い合わせ条件を不完全質問処理によって生成できるので、利用者は全ての可能性を条件として記述する必要はない。この性質を利用すると類似検索も可能である。

本稿では、これらの概念を反映したデータベース応用を記述するためのデータ定義言語と問い合わせ言語を提案する。

2. データベースサービス、既存のデータベース言語、そして不完全質問のための言語

この節では近未来のデータベースサービス産業において、筆者らがMeSODモデルと不完全質問処理の例として扱ってきた鳥類図鑑や大相撲放送の動画データベースをサービスとするような想定をする。この例の中でMeSODモデルと不完全質問処理に基づくデータベース応用の役割について説明する。そして、従来のデータベース言語の役割について見た後、提案する言語の要件についてまとめる。

2.1 近未来のデータベース産業の構図

近未来のデータベースと十分なバンド幅を持つネットワークは図1のように接続され、いつでも家から、出先から必要な情報にアクセスできるようになると考えられている[3]。図1では複数のデータベースと繋がっている地球の形のアイコンが様々な分野の情報提供サービスについてのサービスを表している。ちょうど、電話で言うYellowPageサービスに対応するサービスである。

YellowPageサービスは様々な種類の情報提供サービスを登録しており、利用者からの要求に応じてサービス会社を紹介する。利用者は紹介されたサービス会社に対して問い合わせを投げることによって必要な情報を手にする。

サービス会社はある特定の分野に対する詳細な情報を管理しており、利用者の自宅や出先からのデータに対するアクセスによって収入を得る。この時、大事な点は競合する分野のデータのスキーマ情報は各社の極秘事項であり、YellowPageサービス会社への登録においても完全なスキーマ情報は伝えられない。

しかし、利用者からの問い合わせに対しては利用者の端末に備わっているエージェントソフトとのやり取りによってスキーマ情報を提供できる。

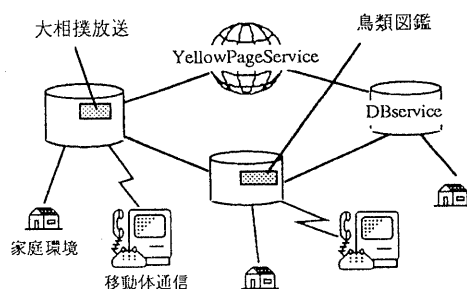


図1 近未来のデータベースサービスとネットワーク

2.2 不完全質問処理の役割

不完全質問処理をベースにした情報提供サービスには大きく分けて二つの市場に以下のような質的な努力をそれぞれ要請する。

(1)情報提供会社の利用者の問い合わせ条件に含まれる不完全さに対する支援

(2)検索エージェント搭載端末が持つ不完全質問の記述支援

(1)の支援は、データベースを中心にした情報提供会社が、できるだけデータをアクセスしてもらうようにスキーマを工夫することで実現される。つまり、索引とすべき項目の決定とそれらの項目に対する書換え知識とを工夫しなければ、他社との競走に負けるからである。この時、不完全質問処理は良い意味でのデータの押売り機構として働く。

(2)の支援は、利用者が購入する検索エージェントソフトが問い合わせ条件に対する不完全さと確かさを汲み取って不完全質問を作ってくれることで実現される。つまり、不完全さに関する制約を利用者の意図を反映するように調整する仕組みを持つような検索エージェントソフトでないと他社との競走に負けるからである。この時、不完全質問処理は不都合なデータに対するフィルターとしての役割を果たす。

2.3 既存のデータベース言語

既存のデータベース言語では対象世界に対する専門的な知識を駆使した複雑な要求を記述したり、そのような問い合わせ結果を使ったデータベース応用の作り易さを重視しているように見える。

(1)関係データベース言語

複雑な条件を宣言的に記述できることがSQLなどの関係問い合わせ言語の利点である[4]。そのような利点が生かされるのは利用者がスキーマ情報とその意味に熟知しており、データ間の複雑な関係についての情報を得ようとするときである。

また、ビジネス応用では指定した属性の属性値の順序関係に基づいて検索結果のリストを表示するようなSQLの問い合わせをホスト言語の中に埋め込んで使われる。

どうせ、ホスト言語に埋め込んで使うなら一つの言語でできた方が良くと言うことが次のオブジェクト指向データベースの利点にもなっている。

(2)オブジェクト指向データベース言語

オブジェクトデータベース言語の利点は、応用を記述する際のデータのモデル化の容易さと言語

の拡張性にある。モデル化の容易さはデータベースの管理単位をオブジェクトと呼ばれる抽象的な単位とし、それに対する更新、削除、挿入にまつわる面倒からプログラムを解放する点に起因する[5]。

また、POSTGRESでは外部関数、演算子の外部定義、ユーザ定義関数の支援によって、従来ながらの問い合わせ言語による問い合わせを様々なデータ構造を持つデータベースでも使えるようにしている[6]。

2.4 新しい言語に対する要件

McSODモデルと不完全質問処理の基本的な仮定は、簡単な問い合わせ条件でも正確には書かれないことである。この仮定は2.1で上げたような例で一般の利用者が使う場合には自然な仮定であろう。

そこで不完全質問処理に基づくシステムを記述するための言語に対する要件を以下のようにおく。

(1)データに対する近さの定義の内包

(2)宣言的な問い合わせ条件

(3)問い合わせ条件の不完全さの記述

(4)問い合わせ条件中のマルチメディアデータのキー指定

(1)の要件は、データ間の近さを計量したり、判定したりする仕組みが不完全質問処理というメタな仕組みからしか使われないということに起因して生じる。つまり、利用者が直接的に距離関数や書換えに関する知識にアクセスできないように、それらを空間という単位の中に押し込める必要がある。

(2)の要件は、不完全質問処理に基づくシステムに対する問い合わせは典型的な条件だけで済むことが特徴であるから宣言的な問い合わせ記述が良い。

(3)の要件は、具体的なデータを指し示して「これに似たデータを求めよ」というような問い合わせを記述するために必要である。具体的なデータを指し示すことから、少なくともそのデータは必ず検索されるので常にその問い合わせ条件は真となる。それでは、不完全質問処理は働かないので問い合わせの中で「これではないかもしれないが」という前置きを付けられるようにすることが、この要件の意図である。

(4)の要件は、マルチメディアデータを直接、問い合わせ条件中の書けるような仕組みを入れることを意味する。

3. データ定義言語

2.4の(1)の要件を満たすために、MeSODモデルにおける空間オブジェクトを定義するための構文を以下のように定める。

空間オブジェクト ::= space { 空間構造定義 [空間関連定義] [距離関数定義] [書換え意味制約定義] [書換え評価基準定義] }

このように空間オブジェクトは最低でも空間構造定義を含まなくてはならない。空間構造定義、空間関連定義、距離関数定義、書換え意味制約定義、書換え評価基準定義は表1のようである。

表1: データ定義構文とその意味

定義構文	意味
structure { 空間構成子 { 空間オブジェクトリスト }}	定義対象の空間がどのような空間を要素にどのような空間構造を持つかを定義
relationship { 関連 関連名 空間オブジェクト }	ある種類のある名前に関連で他の空間オブジェクトの要素にアクセスできることを宣言する。
constraints { 近傍制約リスト, 存在可能性制約リスト }	近傍制約: 定義された空間内の点オブジェクトの守備範囲 存在可能性制約: 削除による書換えを続けるための条件
criteria { 空間オブジェクト間の 半順序関係リスト }	空間オブジェクト毎の距離値の値の大小

空間構造定義、空間関連定義はデータの集合の静的な構造と関連を定義している。距離関数定義、書換え意味制約定義、書換え評価基準定義は書換えに関する知識として持たされる。不完全質問処理のメカニズムがそれらを参照し、候補を探す。

このようにデータ定義の中で書換え意味制約や書換え評価基準を定義しておくことは、間違いに対する許容条件をあらかじめ定義しておくことに外ならない。

3.1 空間構造定義

定義対象の空間構造は、structure節の中でcartesian, quotient, sequence, set構成子によって定義される。これらの空間構成子は空間構造の性質、すなわち距離関数を規定する。したがって、空間構造定義は定義対象の空間内の2点間の距離が他の空間における2点間の距離によって定義されることを表す。図2.3は文献[2]の例で用いた大相撲放送を動画データベースと見なした時の概念図である。

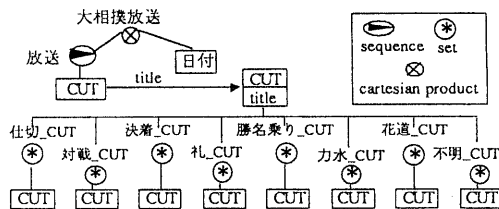


図2 大相撲放送の概念スキーマグラフ

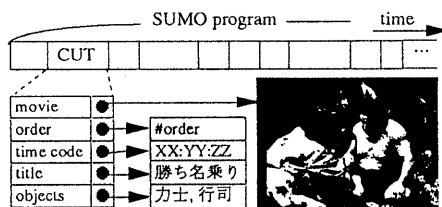


図3 大相撲放送とカットのモデル

また、図4は、これに対する空間構造定義例である。以下に、各空間構成子の使い方を説明する。

(1)cartesian

cartesianに続く {} 内の空間オブジェクトのリストの各空間オブジェクトの直積を空間構造に持つことを表す。その距離は各空間オブジェクトの直積距離をデフォルトとする。例えば、図4(a)の空間構造は空間オブジェクト放送と目付の直積距離空間であることを表している。また、図4(c)の空間構造も直積距離空間で宣言されている。

(2)quotient

quotientに続く二つの空間オブジェクトをX,YとするとXのYによる商を空間構造に持つことを表す。その距離は各空間オブジェクトYの距離をデフォルトとする。図4(d)のCUT/titleはCUTのtitleによる商を空間構造に持つ。直観的には、CUT/titleの点オブジェクトは各タイトル毎のカットの集合に対する名前付のポイントのように振舞う。

(3)sequence

sequence X by X.Yで空間オブジェクトXの要素の系列を要素に持つような空間構造を定義する。X.Yの要素は順序づけられていなくてはならない。例えば、図4(b)はCUT.orderの要素で順序付けされたCUTの系列を要素に持つような空間である。二つの系列間のデフォルトの距離は、相異なる部分系列の最大長である。

(4)set

set XでXの要素からなる集合を要素とするよう

<pre>Space 大相撲放送 { structure { cartesian {日付, 放送} } relationship {} distance { default } constraints { neighbor (distance, 10%) } criteria {} }</pre>	<pre>Space 放送 { structure { sequence CUT by CUT.order } relationship {} distance { default } constraints { neighbor (distance, 10%) } criteria {} }</pre>
---	---

(a)大相撲放送

(b)放送

<pre>Space CUT { structure { cartesian {movie, order, time_code, objects} } relationship { quotient title CUT/title } distance { default } constraints { neighbor (self, 10%) } criteria {} }</pre>	<pre>Space CUT/title { quotient CUT title } relationship { rev_quotient CUT } distance { default } constraints { neighbor (self, 10%) } criteria {} }</pre>
---	---

(c)カット

(d)カットのタイトル
による商

図4 大相撲放送データベースに対するデータ記述例

な空間構造を定義する。すなわち、 X の部分集合を要素とするような空間である。 x_1, x_2 を X の二つの部分集合とすると、その間の距離は、 x_1, x_2 の要素間の距離値の最小値で定義される。

3.2 空間関連定義

空間関連は定義対象の空間オブジェクトから他の空間オブジェクトへの関連を表す。空間関連の直観的な意味は、定義対象の空間オブジェクトの要素や部分集合の近さを別の空間で計量できることの宣言である。空間関連では関数関連、quotient関連、isa関連の3種類の関連をサポートする。

(1)関数関連

定義対象の空間オブジェクトの任意の点オブジェクトが定義された関連によって他の空間オブジェクトのある点オブジェクトに数値的に対応することを意味する。例えば、図5のように空間オブジェクト鳥から名前への関数関連nameが定義されているとする。この時、鳥->nameは空間オブジェクト名前の鳥の名前集合を表す。

この定義では、鳥の近さは名前の近さによって表しても良いことを意味する。

(2)quotient関連

定義対象の空間オブジェクトの任意の点オブジェクトが定義された関連によって他の空間オブジェクトのいずれかの点オブジェクトに分類されることを意味する。図4(c)のCUTの定義中のquotientは全CUTがtitleの要素によって分類されることを宣言している。

また、図4(d)ではrev_quotientという商関連の逆関連によって、その逆の事を表している。つまり、この宣言によって、titleの要素名と'CUT'を '_' で連結した名前で各タイトル毎のカットの集合を表すことができる。

例えば、titleの要素である'仕切'や'対戦'に対して、仕切_CUTや対戦_CUTはそれぞれ'仕切'タイトルと'対戦'タイトルを持つカットの集合を表す。また、これらは共通のカットを持たない。

同様に、図5の空間オブジェクト鳥の定義においても、quotient関連によって全ての鳥はいずれかのシルエットに対応することを表す。つまり、全てのシルエットは少なくとも一種以上の鳥から参照される。また、一つのシルエットはシルエット毎の鳥の集合を指し示すポイントのように扱える。

したがって、シルエット毎の鳥の集合の近さをシルエット自身の近さによって表せるようになる。

(3)isa関連

定義対象の空間オブジェクトが別の空間オブジェクトの特別な場合であることを意味する。したがって、その空間オブジェクトの距離関数はisa関連で示された空間オブジェクトの距離関数によって近さが定義される。例えば、鳥からシルエットへの商関連によってあるシルエットに属する鳥だ

```
Space 鳥 {
  structure {
    cartesian {色パターン, 鳴き声}
  }
  relationships {
    functional name 名前
    quotient シルエット シルエットパターン
  }
  distance { default }
  constraints {
    neighbor(色パターン, 5%)
    neighbor(鳴き声, 3%)
    exist (鳥->name) -> exist(鳥)
  }
  criteria {
    シルエットパターン > 色パターン
    シルエットパターン > 鳴き声
  }
}
```

図5 空間オブジェクト鳥の定義例

けを集めた空間オブジェクトを定義できるが、この空間オブジェクトは鳥に対してisa関連を持つ。したがって、あるシルエットに属する鳥の近さの定義は鳥の近さの定義を示す。

3.3 距離関数定義

空間構造定義における空間構成子によって距離関数は定まる。このデフォルトの距離関数以外に特別な距離関数を定義できる。定義文は、distanceに続く{}の中にデータ操作と組み込みの距離関数を使って記述する。

例えば、色パターン検索の場合の距離関数定義は以下のとおりである。まず、ある色パターンに対する間違いパターンを問い合わせによって求める。求められた結果ともとの色パターンを含めた色パターン集合と与えられた色パターンのユークリッド距離の最小値を距離値と定義する[7]。なお、色パターンは色のCIE-Luv表現を要素とするような12次元のベクトルで表現されている。

また、図4のdefaultは空間構成子によって定義される距離関数を表す。

3.4 書換え意味制約定義

複雑な空間構造を持つ空間オブジェクトでは、構成要素の空間オブジェクトについての制約を与えることができる。制約はconstraintsに続く{}の中で以下の2種類の制約のリストを記述する。

(1)近傍制約

近傍制約の定義構文は以下のとおりである。

近傍制約 ::= neighbor(空間オブジェクト, 範囲指定)

範囲指定 ::= 数値% ; 数値# ; 数値 ∈ [0,1]

この記述は、指定した空間オブジェクトにおいては範囲指定によって示される範囲内の値にしか書換えられないことを表す。範囲指定において、数値%は空間オブジェクト内の全ての要素数を与えられたデータとの距離の小さい順に整列した時の上位数値%を表す。数値#は、距離の小さい順に整列した時の上位数値位を表す。[0,1]の間の数値は、距離値を表す。

例えば、図5においては色パターンの空間では上位5%以内の違いであれば許容できることを、鳴き声の空間では上位3位以内の違いであれば許容できることを定義している。

空間構成子によって複雑な構造を指定した空間オブジェクトにおいては、空間構造の定義中に現れる空間オブジェクトの近傍制約を上書きできる。例えば、整数の距離空間の近傍制約が定義済みと

する。複合空間オブジェクトで整数の空間オブジェクトを含む場合には、近傍制約の定義を上書きできる。

また、図4の近傍制約中のselfは定義対象の空間オブジェクトを参照する変数である。

(2)存在可能性制約

存在可能性制約は以下のように記述する。

存在可能性制約 ::= exist(空間オブジェクト) → exist(空間オブジェクト)

ある空間オブジェクトの書換えが他の空間オブジェクトに対する存在可能性を決定する場合に記述する。制約記述に使える空間オブジェクトは、空間構造定義か空間関連定義中に現れる空間オブジェクトと定義対象の空間オブジェクトのいずれかである。

図5の存在可能性制約は、鳥の名前を指定した条件との論理積で表された問い合わせ条件が名前空間の近傍制約に違反するような書換えが必要とすれば、他の条件の如何に関わらず目的の鳥は存在しないと結論する。

この制約が書かれていない場合には、問い合わせ条件中の鳥の名前以外の項目に関する条件に対する書換えの成功によって候補を検索する。この例では、名前の条件との論理積である問い合わせ条件では、名前を間違えることは鳥の指定に致命的であるという仮定をしている。

(3)点オブジェクト毎の制約

書換え制約に関しては点オブジェクト毎に定義することができる。その例は、鳴き声検索の例がある[7]。

3.5 書換え評価基準定義

書換え評価基準は、criteriaに続く{}の中で以下のように記述する。

書換え評価基準 ::= 空間オブジェクト > 空間オブジェクト

記述中に現れる空間オブジェクトの書換えコストのどちらがコストが高いかを表す。例えば、図5のcriteriaでは、指定された条件に対する書換えについてシルエットについての書換えコストが色パターン、鳴き声についての書換えコストよりも高いという仮定を置いている。

したがって、このような定義に基づく鳥類データベースでは、書換えコストが安くなるような鳥は指定したシルエットに類似しているシルエットを持つ鳥の集合の中で、色パターンや鳴き声がそれぞれ類似しているような鳥である。

4. 問い合わせ言語

2.4の(2),(3)の要件に対応してSQL風の問い合わせ構文を基本に、incomplete節を提案する。問い合わせ文にincomplete節を付けられるようにすることで、問い合わせ条件自身の不完全さに対する言明を記述できるようにする。その結果、応用プログラムで不完全質問処理をコントロールできる。

4.1 問い合わせ式

問い合わせ式は以下の通りである。

問い合わせ式 ::= 問い合わせ指定 [不完全質問指定]

問い合わせ指定 ::= select 空間オブジェクト参照
(探索条件) 空間オブジェクト参照

4.2 探索条件

一般的な探索条件に加えて、論理式の正規表現を可能にすることによって系列データの中にある属性の変化の様子を記述できる[2]。

空間オブジェクトの指定方法には、ドット記法によって空間構造における特定の空間オブジェクトを指定する方法がある。例えば、鳥の例では鳥色パターン=getColorPattern()は、与えられた色パターンを持つような鳥に対する条件を表す。

また、空間オブジェクトの関連を辿って空間オブジェクトを指定する方法は、->記法を使う。鳥の例では、鳥->シルエット=getSilhouette()は、与えられたシルエットパターンを持つような鳥を求める際の条件として使われる。なお、getColorPattern()やgetSilhouette()を外部関数と呼ぶが、詳細は5.で説明する。

4.3 不完全質問指定

問い合わせ指定に続くincomplete節によって、不完全質問処理機構に渡すべきパラメータを記述できる。その結果、問い合わせに対する利用者の自信や候補としての最低条件を記述できる。簡単な例には、鳥類図鑑におけるシルエットのガイダンス機能がある。この例では、incomplete節を空とすることによって指定した問い合わせ条件に対する不完全質問処理の結果として得られる全ての候補を要求している。

また、incomplete節を付けないことは、完全なマッチングに基づく検索を指示したことに対応する。incomplete節の中にはデータ定義に現れる書換え制約、書換え評価基準、距離関数を記述できる。これらが有効なのはその問い合わせ中だけである。

```
select sequence 取組 CUT by CUT.order {
  (CUT.title='仕切')+ (CUT.title='対戦')+
  (CUT.title='決着')+ (CUT.title='礼')+
  (CUT.title='勝名乗り')+ (CUT.title='力水')+
} 放送 incomplete {
  constraints {
    neighbor(Bseq, ε);
    exist(対戦_CUT) → exist(取組)
  }
  criteria {
    Bseq > 仕切_CUT > 対戦_CUT > 決着_CUT > 礼_CUT
    > 勝名乗り_CUT > 力水_CUT > 花道_CUT
  }
}
```

図6 取組ダイジェスト要求に対する不完全質問文

(1)constraints

空間構造定義に現れる空間オブジェクトリストのメンバーに対する近傍制約と存在可能性制約のリストを記述する。

図6の近傍制約は系列に対する問い合わせ条件の真偽値の系列 B_{seq} に、連続して現れる偽の数によって、問い合わせ条件の書換えを制限する。

図6の存在可能性制約は、書換えによって得られる取組の候補系列に対戦_CUTに属するカットがあれば、その候補系列が問い合わせで要求する取組である可能性が有ることを示している。したがって、書換えによって得られる取組の候補系列には対戦_CUTに属するカットが存在することが保証される。

(2)criteria

図6の書換え評価基準は、以下の二つの性質を持つような系列を得られるような評価基準である。系列に対する問い合わせ条件の真偽値の系列 B_{seq} の中で真となる部分の系列ができるだけ長くなるような系列である。また、一部のタイトルの系列が抜けていたり、指定したタイトル以外の系列が現れても仕切、対戦、決着、礼、勝名乗り、力水、花道の順にそれぞれのタイトルに対応する部分系列が現れるような系列である。

(3)距離関数操作

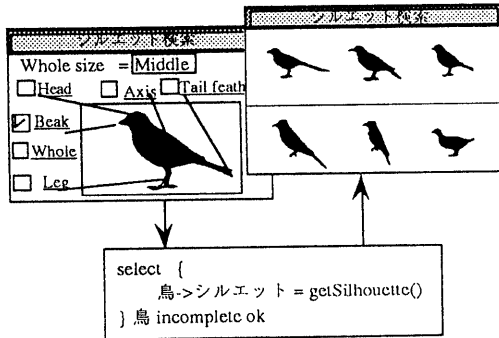
空間オブジェクトの距離関数に対する拡大縮小値を与えられる。例えば、シルエット検索の例では利用者が選択したシルエットの各部位毎に拡大縮小値を与えられる[7]。

4.4 対話検索オプション

Incomplete ok { }によって、検索結果に対する利用者や応用からの神託を受けられる。結果として得られた候補に対して、真偽を与えることで問い

合わせ処理の終了/続行を決められる。

okオプション付の不完全質問の結果の集合resultsに対する神託が偽であった場合、constraints句にnot resultsが加わえたものを新しいincomplete節とする。問い合わせ式に新しいincomplete節を付けた不完全質問の結果が空になるか神託が真となるまで不完全質問処理は繰り返される。



指定したシルエットに似たシルエットを求めよ

図7 シルエット検索と不完全質問式のokオプション

図7のシルエット検索は、この例である。つまり、利用者が指定することで与えられるシルエットパターンは実際にデータベースに存在するから、問い合わせ条件は必ず真である。しかし、incomplete節が付いているので候補を返す。これを対話的に繰り返し、候補を絞り込むには、okオプションを付ける。

5. 外部拡張に関する機能

2.4の(4)の要件に対して、4.で示した構文においてマルチメディアデータをキーにした問い合わせ条件を記述する方法と、そのような問い合わせ条件を不完全質問処理によって扱う方法について論じる。

(1)問い合わせ条件中のマルチメディアデータ

マルチメディアデータは定数と考えられるから、問い合わせ条件中に現れても不思議ではない。そこで、利用者定義の外部関数とポイントによってマルチメディアデータの入力を支援する。例えば、図8の端末で鳴き真似を問い合わせ条件に含める場合には外部関数getVoiceを検索Agentで定義して、利用者からの問い合わせ記述を鳥.鳴き声 = getVoice()とする。

ネットワークにバケットとして出力する際には、getVoiceをポイントとしてバケット中の音声データの先頭位置に対するオフセットを渡す。

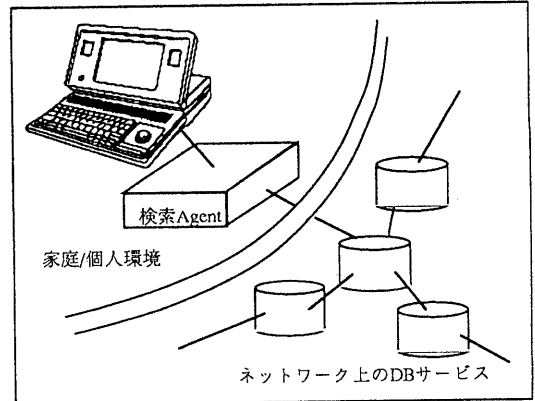


図8 ネットワーク上のDBサービスと個人環境

(2)マルチメディアデータから索引への分解

類似検索するためにはマルチメディアデータから索引項目毎の値を取り出せなくてはならない。

空間オブジェクトの定義では、空間構造に現れる空間オブジェクトの点オブジェクトにマルチメディアデータを対応付けられなくてはならない。

そこで比較演算子の定義の中で呼び出されるextract関数という特別な外部関数を空間オブジェクト毎に定義できるようにする。

例えば、鳴き真似を入力として鳴き声の類似検索をする際は、鳥.鳴き声 = getVoice()という条件がシステムに与えられる。これに対して、空間オブジェクトの鳴き声の外部関数として定義されたextract関数が、鳴き声空間の新しい点オブジェクトを生成して各特徴に値を格納する。その結果、図9のように与えられた問い合わせ条件はより詳細な特徴に関する条件の論理積に変換される。

```

鳥. 鳴き声 = getVoice()
⇔ 鳥. 鳴き声 = extract(getVoice())
⇔ 鳥. 鳴き声. 特徴1 = extract(getVoice()).特徴1 and
   鳥. 鳴き声. 特徴2 = extract(getVoice()).特徴2 and
   ...
   鳥. 鳴き声. 特徴7 = extract(getVoice()).特徴7
    
```

図9 マルチメディアキーの分解による問い合わせ変換

6. むすび

本稿では、提案中のMeSODモデルと不完全質問処理に基づくデータベース応用を記述するためのデータ定義および問い合わせ言語を提案した。MeSODモデルと不完全質問処理の前提では、問い合わせ条件の意味論的な複雑さ以上に利用者の指定

する条件の不正確さに重きを置いている。そこで、データ定義において予測可能な不完全さを記述できるようなデータ定義構文と、問い合わせ条件自身に対する不完全さの言明を利用者にできるような問い合わせ構文を提案した。また、マルチメディアデータを問い合わせの中に含まれるような言語機能についても提案した。

文献

- [1]田渕仁浩,村岡洋一:"MeSODモデルにおけるオブジェクト概念",信学論D-1, J74-D-1, 4 (Apr.1991).
- [2]田渕仁浩,村岡洋一:"動画像データベースにおける不完全質問とMeSODモデル",情処学研報, 92-DBS-89-2, pp.9-18 (1992).
- [3]日経サイエンス 特集コンピュータネットワーク,11月号, 1991
- [4]Date C. J.: "A Guide to THE SQL STANDARD", Addison-Wesley,1987. 邦訳(芝野耕司, 岸本令子), "標準SQL", 凸版 (1988).
- [5]Deux et al.: "THE O₂ SYSTEM", Comm. ACM,34, 10, pp.34-48 (Oct. 1991).
- [6]Stonebraker M., Kemnitz G.: "THE POSTGRES NEXT GENERATION DATABASE MANAGEMENT SYSTEM", Comm. ACM,34, 10, pp78-92 (Oct. 1991).
- [7]田渕仁浩,村岡洋一:"マルチメディアデータベースと認識的な問い合わせ処理",人工知能学会SIG-HICG-9202-3 (1992).