

仮想世界データベースシステムにおけるマルチモーダル問合せ言語の開発

Development of a Multi-modal Query Language for the Virtual World Database System

矢野 ナホコ[▼] 横川 明子[▼]
増永 良文[▲]

Nahoko YANO Akiko YOKOGAWA
Yoshifumi MASUNAGA

我々はVRシステムに特有な機能をデータベースに取り入れた新世代データベースシステムとして、仮想世界データベースシステム (Virtual World Database System: VWDB) の構築を進めてきた。既に、プロトタイプが実装され、その上で共同作業に向けたトランザクション管理などの研究が行われ、仮想世界をデータベースとする VWDB の概念が具体化されてきている。本研究では、VWDB のさらなる機能拡張に向けて、3次元データベース空間におけるマルチモーダル (音声とマウスによる入力) 問合せ言語を開発する。ユーザからの音声入力をルール認識により問合せ言語に変換することで、SQLのような問合せ言語を意識することなく、より実世界に近い自然なインタラクションで VWDB への問合せが実現される。

The Virtual World Database System (VWDB) is currently under development at Ochanomizu University. The VWDB is a novel database system in that it has a variety of VR system features in addition to traditional database functions. It provides a 3D database space for users, and a prototype has been implemented to provide a transaction management capability for shared work environment. This paper reports a new VWDB function which is a VWDB multi-modal query language. It uses a mouse and voice for interaction in the 3D database space so that users can issue a variety of queries as if they were in the real world.

1. はじめに

我々は、VRシステムに特有な機能をデータベースに取り入れた新世代データベースシステムとして、仮想世界データベースシステムの構築を進めてきた。これまで、VWDBデータベーススキーマ言語の定義[1]や共同作業に向けたトランザクション管理の研究[2]が行われ、仮想世界をデータベースと

[▼] 学生会員 お茶の水女子大学大学院人間文化研究科博士前期課程 {nahoko, acco}@db.is.ocha.ac.jp

[▲] 正会員 お茶の水女子大学理学部情報科学科 masunaga@is.ocha.ac.jp

するVWDBの概念が具体化されてきている。しかしながら、肝心の問合せ機能のモデル化と実装は行われていない。そこで本研究では、VWDBのさらなる機能拡張に向けて、3次元データベース空間におけるマルチモーダル (音声とマウスによる入力) 問合せ言語を開発する。ユーザからの音声入力をルール認識により問合せ言語に変換する。ルール認識とは、事前に語彙辞書に登録したセンテンスのみを音声エンジンが認識する方式であり、問合せのキーワードとなる単語を的確に把握し、効率的に処理を行う。これによりユーザはSQLのような問合せ言語を意識することなく、より実世界に近い自然なインタラクションでVWDBへの問合せが行える。

2. VWDB システム

2.1 システム構成

VWDBはバックエンドのデータベースシステムと、フロントエンドのVRシステム群で構成される。フロントエンドのVRシステムはNVR (Networked Virtual Reality System) により互いに接続されており、複数のユーザが同一の仮想環境を共有できるようになっている。

2.2 VWDB スキーマ

VWDB では仮想世界および仮想世界オブジェクトの生成をこれまでのデータベースと同様にスキーマレベルで行う。仮想世界オブジェクトは、3次元空間における位置、姿勢の空間的属性、名前や所有者などの非空間的属性と、形状属性を持つ。VWDB では、仮想世界オブジェクトのスキーマを「カテゴリ」としカテゴリはクラス概念を継承する。全てのカテゴリに継承されるルートカテゴリを「VWDB カテゴリ」とする。VWDB カテゴリでは、全ての仮想世界オブジェクトが持つ識別子と位置属性や姿勢属性が定義されており、全てのカテゴリはVWDBカテゴリを継承する。

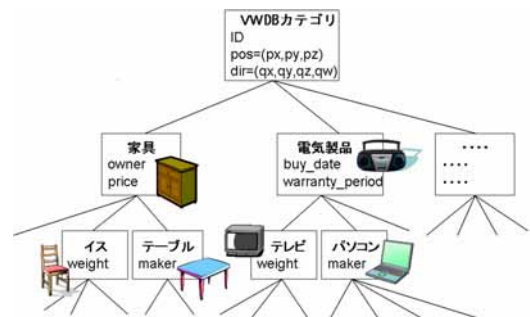


図1 VWDBスキーマ

Fig.1 VWDB Schema

図1にカテゴリ階層の一例を示す。例えば、イスカテゴリの属性は、スーパーカテゴリから継承した(ID, pos, dir, owner, price)に(weight)を加えた(ID, pos, dir, owner, price, weight)と図1に示される形状属性となる。このように関係づけられたカテゴリ階層が、VWDBデータベースのスキーマとなる。仮想世界オブジェクトはカテゴリのインスタンスとして生成されるので、「イス」や「テーブル」といったセマンティクスを持ちあわせる。

3. VWDB における空間関係

実世界さながらのVWDB空間で、ユーザは目の前に存在するオブジェクトに様々な問合せを発行したいであろう。我々はVWDBにおける問合せとして「所有者がAさんのものは？」といった非空間的属性に関する問合せと、「机の上

にあるものは何ですか？」といった空間関係に関する問合せを想定する。非空間的の属性に関する問合せを処理するには、オブジェクトを格納しているカテゴリリレーションを使用すればよい。一方、空間関係に関する問合せを処理するには新たなリレーションを用意する必要がある。そこで空間関係の問合せを行う前提として、オブジェクト間の空間関係をどのように表現し格納・保持するかについて論じる。

3.1 空間関係の表現

VWDB 空間は 3 次元空間であるので、オブジェクト間の空間関係に関する様々な問合せが想定される。我々は、VWDB における空間関係の問合せとして、空間関係の 3 大要素 [3] として挙げられる距離・方向・位相と、さらにアバタからの可視性を判断する 4 種類の問合せを想定する。空間関係は以下のように 3 項組または 4 項組の述語で表現する。

距離 : Distance(O_A, O_B, θ, dis)

方向 : Direction(O_A, O_B, dir)

位相 : Topology(O_A, O_B, top)

可視性 : Visibility (A, O_B, vis)

ここに O_A : 基点オブジェクト, O_B : 対象オブジェクト, A : アバタ, θ : 比較演算子 ($=, <, >, \leq, \geq$), dis : 距離 (数値), dir : $\overline{O_A O_B}$, top : 位相 (*meet, disjoint, inside, contains*), vis : *true / false* .

3.2 空間関係の格納

VWDB で空間関係を問合せする前提として、ある時刻における仮想世界の状態をあらかじめデータベースに格納しておく必要がある。まず Distance(O_A, O_B, θ, dis) の格納について考える。 θ が " $=$ " 以外の場合、 dis は無限に存在し格納不可能である。例えば O_A と O_B の距離を a とし Distance($O_A, O_B, <, dis$) の格納を考えた場合の dis は、 $dis = \{ x | x \in (a, \infty) \}$ で無限個である。そこで Distance($O_A, O_B, =, dis$) を満たす Distance(O_A, O_B, dis) を Distance リレーションに格納する。方向関係、位相関係については 3 項組をそのまま格納するものとし、空間関係リレーションのスキーマは下記の通りである。

距離 : Distance(O_A, O_B, dis)

方向 : Direction(O_A, O_B, dir)

位相 : Topology(O_A, O_B, top)

なお、VWDB ではアバタは頻繁に位置を変更、向きも変えるものと仮定し、アバタの空間的情報は格納しないので可視性に関するリレーションは用意しない。仮想世界上で、オブジェクトが生成・移動・削除された場合は、トリガ機能で空間関係リレーションも更新し空間関係を保持する。

4. VWDB におけるマルチモーダル問合せ言語の設計

VWDB では、述語論理をもとにした体系で問合せを表現し、それを満足する集合を問合せ結果とする。

4.1 空間関係の問合せ

空間関係の問合せ述語として、空間関係の 3 大要素 [3] として挙げられる距離・方向・位相と、さらにアバタからの可視性を判断する問合せを定義する。なお、以下に示す述語の中で、自分のアバタは基点オブジェクトの一つとして扱い、他人のアバタは、基点オブジェクトあるいは対象オブジェクトの一つとして扱う。ユーザは自分の視点から問合せを行うであろうと想定し、自分のアバタを対象オブジェクトとはしない。

距離に関するプリミティブな問合せとは、述語 Distance のみを使って定義される問合せとすると、15 種類が考えられ

る。しかし、この中で結果が無限に存在する問合せを除去した次の 10 個が距離に関する「意味のあるプリミティブな問合せ」(これを MPQ: meaningful primitive queries といおう) とする。

(dis-1) Distance($O_A, O_B, =, ?$)

(dis-2) Distance($O_A, O_B, ?, d$)

(dis-3) Distance($O_A, ?, \theta, d$)

(dis-4) Distance($?, O_B, \theta, d$)

(dis-5) Distance($O_A, ?, =, ?$)

(dis-6) Distance($?, O_B, =, ?$)

(dis-7) Distance($O_A, ?, ?, d$)

(dis-8) Distance($?, O_B, ?, d$)

(dis-9) Distance($?, ?, \theta, d$)

(dis-10) Distance($?, ?, =, ?$)

同様に、方向・位相・可視性に関する MPQ はそれぞれ 7 種である。なお、問合せにおける方向パラメタは、絶対座標 (ワールド座標) を基準にした東/西/南/北/北東/北西/南東/南西/天頂/天底と、アバタの視点 (ローカル座標) を基準にした前/後/左/右/右前/左前/右後/左後/上/下のうちのひとつとする。

問合せ例 1) O_A と O_B の距離は?

Distance ($O_A, O_B, =, ?$)

問合せ例 2) 私の前にあるものは?

Direction ($A, ?, F$)

問合せ例 3) O_A の中にあるものは?

Topology ($O_A, ?, contains$)

以上、Distance, Direction, Topology, Visibility の 31 の問合せを、VWDB の空間関係の MPQ と定義する。

4.2 非空間的属性の問合せ

オブジェクトのカテゴリを A とし、 A の非空間的属性のスキーマを $A (A_1, A_2, \dots, A_n)$ とする。 A_1, A_2, \dots, A_n に対して条件指定したい属性には条件を、所望する属性には "?" を、それ以外の属性には "-" を指定する。これを非空間的属性の問合せ述語とする。

問合せ例 4) 所有者が太郎さんの机の値段は?

Table ("Taro", ?, -)

例 4 では机カテゴリのスキーマは Table(owner, price, maker) とする。このように 1 つのカテゴリ名 (Table) と、0 個以上の条件 (owner="Taro") および 0 個以上の属性値名 (price=?) で表現される問合せを VWDB の非空間的属性の単純質問とする。

5. マルチモーダル問合せ処理

5.1 音声認識方式

VWDB のマルチモーダル問合せは、通常のディスプレイに表示された VWDB 空間を視ながら音声入力とマウス入力で行う。オブジェクトの指定は音声入力と同期したマウス操作 (ピッキング) で行い、音声認識はルール認識方式で行う。ルール認識方式とは事前に語彙辞書に登録したセンテンスのみを音声エンジンが認識する方法であり、発話内容の意味を的確にとらえることができ、音声コマンドによる制御や効率的な情報検索に向いている。VWDB における問合せを考えた場合、使用されるキーワードが限定されるため、ディクテーション認識による複雑な自然言語処理を行うより、ルール認識の方が確かつ効率的に処理できる。

図 2 はマルチモーダル問合せ処理系を示す。音声入力データとピッキングデータがマルチモーダル入力認識モジュールで処理される。音声入力データがあらかじめ定義したルールに一致した場合のみ音声認識結果として認識されたルール名などが構文解析部にわたる。構文解析部では音声認識結

果に基づき、ピッキングデータとの照合やアバタの情報を取得し、問合せ言語へ変換する。さらにメッセージ生成部で問合せ言語から SQL 文へ生成し発行する。

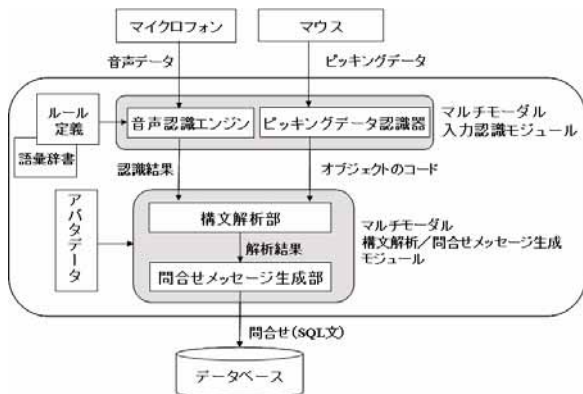


図2 マルチモーダル問合せ処理系

Fig.2 Multimodal Query Processing System

例えば、ユーザから「これとあれの距離はいくらですか?」という問合せがあった場合は、構文解析部でこの音声とマウスによる入力を $Distance(O_1, O_2, =, ?)$ という問合せ言語に変換する。さらに問合せメッセージ生成部で「SELECT dis FROM Distance WHERE $O_A=O_1$ AND $O_B=O_2$ 」というリレーション $Distance(O_A, O_B, dis)$ へのSQLに変換してデータベースに問合せを発行する。空間関係のMPQも今日想定する非空間的属性の問合せも単純質問であり、問合せ言語からSQLへの変換はほとんど自明であるので、本論文では音声データから問合せ言語への変換に焦点をあてて述べる。

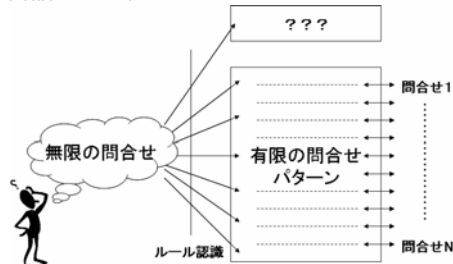


図3 ルール認識による問合せのパターン化

Fig.3 Making Queries into Patterns by Rule Recognition

図3に示すように、ユーザからの音声による無限の問合せ表現を31のMPQまたは非空間的属性の問合せに変換するために、ルール認識方式では的確なルール設計が必須となる。

5.2 空間関係の問合せのためのルール設計

空間関係の問合せでは、基点オブジェクトと対象オブジェクトの区別は重要である。例えば「これの東にあるものは?」という問合せと「これが東にあるものは?」という問合せについて考える。前者は「これ」の東側に存在するオブジェクトが検索され、後者は「これ」の西側に存在するオブジェクトが検索されることになり結果は大きく異なる。我々は、この区別を正確に判断するために、基点オブジェクトは「これ/この人」、対象オブジェクトは「あれ/あの人」と発声すると共にマウスでクリックしてもらうこととする。ただし自分のアバタを指し示すことはできないので、代わりに「私/自分」と発声してもらう。このような前提条件を定めた上で、空間関係の問合せを特徴付ける単語を、ルールとして次のように定義する。

<dis> = (<INTEGER> | <FLOAT>);
 <dir> = (前 | 後 | 左 | 右 | 上 | 下 | 右前 | 左前 | 右後 | 左後 | 東 | 西 | 南 | 北 | 北東 | 北西 | 南東 | 南西 | 天頂 | 天底);
 <top> = (meet | disjoint | contains | inside);
 <θ> = (より小さい | より大きい | 以下 | 以上 | でない | である);
 <vis> = (見える | 見えない);
 <obj1> = (これ | この人 | 自分 | 私);
 <obj2> = (あれ | あの人);
 <avt> = (myname | username1 | username2 | ...);
 <DIS> = 距離; <DIR> = 方向;
 <TOP> = 位相; <VIS> = 可視;

この表記は、拡張BNF構文 (ABNF) の一種であるJSpeech Grammar Format (JAGF) に従っている。<>はルール名を表し、| は“または”を表す。<INTEGER><FLOAT>については、整数・小数を認識するルールとする。各MPQに固有な非終端記号の集合をRとする。例えば、 $(dis-1) Distance(O_A, O_B, =, ?)$ 「これとあれの距離は?」の場合、「これ」に対応したobj1と「あれ」に対応したobj2, さらに「距離」に対応したDISの3つの要素を持つ集合がRとなり、 $R_{dis-1} = \{obj1, obj2, DIS\}$ となる。各Rは互いに同一とならないようにルール設計するので、31のMPQへの対応付けが可能となる。

このRをもとに、さらに語順、助詞などを考慮して、31のMPQに対応したルールを定義する。例えば、 $(dis-1) Distance(O_A, O_B, =, ?)$ のルール設計について考える。 $R_{dis-1} = \{obj1, obj2, DIS\}$ であるので語順としては6通りが考えられるが、このうち自然言語としてもっともらしい2通りの語順を考慮してルールを定義する。

```
public <dis-1>=<obj1><obj2> | <obj2><obj1> <DIS><stop>;
<obj1>=<これ | この人 | 私 | 自分> <tofrom>;
<obj2>=<あれ | あの人> <tofrom>;
<DIS>= 距離 [関係][が][は];
<tofrom>= [から][まで][との][と][は][の][みて][を]
[基準][と][して][が];
<stop>=[は][どれ][どの][どう][いくら][何][メートル][くらい][に][です][なり][なる][ます][か][を][教え][見つけ][探して][探せ][表示][示し][示せ][せよ][し][る][て][くれ][ます][か][ちようだい][下さい];
```

このルール定義<dis-1>により認識される文の例は「これあれ 距離」「私からあれまでの距離はいくらですか」「あの人と自分の距離を教えてください」などである。このように予想される助詞を"[]"(その語句があってもなくてもよい)により列挙することで、ある程度多様な表現に対応できる。また"|"により語順の転置にも柔軟に対応する。他の30のMPQに対しても同様の方針でルール定義を行う。

5.3 非空間的属性の問合せのためのルール設計

非空間的属性の問合せを特徴付けるルールとして、全てのカテゴリ名、属性値名、属性値を定義する。

```
<category> = (家具 | ベッド | 棚 | 机 | イス | ...);
[カテゴリ][に][所属する][属する][もの][の中で];
<attributeS> = (所有者 | 製造地 | ...);
<attributeN> = (価格 | 値段 | 重さ | ...);
<value> = (太郎 | 花子 | 日本 | イタリア | ...);
<condition> = (<attributeS> が<value>
| <value> が<attributeS> | <attributeN> が<num><θ>
| <num><θ> [の] <attributeN> [で][の][である];
<num> = (<INTEGER> | <FLOAT>) <tani>;
```



```
<tani> = [円][キロ][センチ][ミリ][メートル][グラム]...;
<stop> = [の][である][もの][オブジェクト][は][どれ][どこ][に][あり][ます][です][か][を][教え][見つけ][探して][探せ][表示][示し][示せ][せよ][し][る][て][くれ][ます][か][ください][下さい];
```

「製造地がイタリア」や「5万円以内の価格」というように属性値名と属性値が組になると<condition>として認識される。非空間的属性の問合せにおける非終端記号の集合 R は、1つのcategoryと0個以上のconditionおよびattributeである。この条件をもとに非空間的属性の問合せルールを次のように定義する。

```
public <non-Spatial> = <condition>* <category>
<condition>* [( <attribute>[と] )*] <stop>;
```

“*”は0回以上の繰り返しを意味する。このルール定義<non-Spatial>により認識される文の例は「イス」「机の価格」「家具の中で所有者が太郎のもの」「値段が3千円以上の家具の製造地を教えてください」などである。

6. 実装

本研究では、ルール認識エンジンにアドバンスド・メディア社の AmiVoice を用いて実装を行った。

6.1 マウスデータと音声データの統合

VWDB 空間では問合せとは関係なく常にユーザのマウスイベントを監視しており、任意のオブジェクトをクリックした場合、対象オブジェクトのコード (ID) を時間順にバッファする。音声認識結果の中に、これ/この人/あれ/あの人が含まれていた場合は、その個数によって次のように処理する。

(a) 1つ含まれていた場合

最新のピッキングデータを対応させる。

(b) 2つ含まれていた場合

2つの語彙が発声された時間順序に合わせて、最新のピッキングデータとその前のピッキングデータを対応させる。

6.2 語彙の登録機能

音声で問合せを行う前提として、新しい属性値が入力された場合には、音声認識エンジンやルール定義に追加する必要がある。我々はこれをアプリケーション上から直接行えるようにした。新たな属性値の表記 (漢字) とよみ (ひらがな) をウィンドウ上で入力してもらい、これを認識エンジンの語彙辞書とルール<value>へ自動的に追加する。新規登録された語彙は、即時認識可能となり、問合せにも反映される。

6.3 問合せ方法

問合せは、通常のディスプレイに表示された VWDB 空間を視ながら行う。ヘッドセットを使用し音声の入力を行う。ユーザからの音声入力がいずれかのルールに一致した場合には、認識した文章を表示し、問合せ処理を行い、結果を表示する。結果の表示方法は、オブジェクトの強調表示、オブジェクトに付加した3Dテキストによる表示、および問合せウィンドウにテキスト表示で行う。

実際にオブジェクトを30ほど配置した VWDB 空間で問合せを行った結果例を図4に示す。アバタは現在、空間の中心付近にいる。は、「これとあれの距離は何メートルですか」という音声入力とともに、机を2つクリックした場合の結果である。認識された文章が、右上の問合せウィンドウに表示され、問合せ処理が行われる。結果として机が強調表示され点滅するとともに、3Dテキストと右上の問合せウィンドウに「2.23m」と表示される。

上記の例で示すとおり、本論文で提案したモデルにより自

然言語に近い発話で VWDB の非空間的属性および空間関係の問合せが実現することが確認された。



図4 問合せ結果例

Fig.4 A Sample Query Result

7. まとめと今後の課題

VWDB において音声とマウスを使用したマルチモーダルな問合せを行うために、まず空間関係の表現方法と格納法について論じた。次に、述語論理をもとにしてマルチモーダル問合せ言語の設計を行った。さらに、マルチモーダル問合せ処理系を示し、ルール認識により音声入力を問合せ言語に変換するためのルール設計について論じた。音声認識エンジン AmiVoice を用いて実装を行い、その結果を示した。

今後の課題は、透明度やワイヤフレーム表示を利用し VR の特性を活かした表示法の検討を行うとともに、より複雑な問合せにも対応できるように拡張していくことである。

【文献】

- [1] 渡辺知恵美, “仮想世界データベースシステムの設計と実装,” お茶の水女子大学学位論文, 2003.
- [2] 渡辺知恵美, 大杉あゆみ, 佐藤こず恵, 増永良文, “仮想世界データベースシステムにおける共有型作業環境のためのトランザクション概念の導入,” 情報処理学会論文誌: データベース (TOD), Vol.43, No.SIG9 (TOD15), pp.55-67, 2002.
- [3] M. Egenhofer, “Spatial Relation: Models, Interfaces, and their Future Application,” Proceeding of Advanced Database Symposium '96, Vol.8, No.4, pp.403-424, 1996.

矢野 ナホコ Nahoko YANO

2005 お茶の水女子大学大学院人間文化研究科博士前期課程数理・情報科学専攻修了。仮想世界データベースシステムの研究・開発に従事。日本データベース学会学生会員。

横川 明子 Akiko YOKOGAWA

お茶の水女子大学大学院人間文化研究科博士前期課程数理・情報科学専攻在学中。仮想世界データベースシステムの研究・開発に従事。日本データベース学会学生会員。

増永 良文 Yoshifumi MASUNAGA

お茶の水女子大学理学部情報科学科教授。1970 東北大学大学院工学研究科博士課程修了, 工学博士。データベースシステムの研究・開発に従事。情報処理学会および電子情報通信学会フェロー。日本データベース学会会長。著書に「リレーショナルデータベース入門 [新訂版]」(サイエンス社)など。