仮想世界データベースシステムにおける マルチモーダル融合問合せ処理

横川　明子†　　増永　良文††

† 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
†† お茶の水女子大学 理学部情報科学科
〒 112–8610 東京都文京区大塚 2-1-1
E-mail: †akiko@db.is.ocha.ac.jp, ††masunaga@is.ocha.ac.jp

あらまし　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，VWDB における問合せ言語として，マルチモーダル問合せ言語の基礎開発がなされ，ユーザは音声とマウス入力によって，SQL のような問合せ言語を意識することなく，実世界に近い自然なインタラクションで VWDB への単純な問合せを行うことが可能となっている．しかしながら，この言語は仮想世界における距離・方向・位相・可視性に関して，個別に問合せを発行できるものの，それらを融合した問合せを発行することはできなかった．そこで我々は，距離・方向・位相・可視性に関する問合せを組み合わせることが可能な融合文法を導入したが，本報告ではそのうち二つの問合せを組合せた質問を処理できる VWDB での複雑なマルチモーダル問合せ処理系の開発したので報告する．

キーワード　VWDB，3 次元空間データベース，バーチャルリアリティ，マルチモーダル，問合せ言語，シグネチャ

Multimodal Fusion Query Processing in the Virtual World Database System

Akiko YOKOGAWA† and Yoshihumi MASUNAGA††

† Graduate Division of Mathematics and Computer Science(Master's Program), Ochanomizu University
†† Department of Information Science, Faculty of Science, Ochanomizu University
2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610 Japan
E-mail: †akiko@db.is.ocha.ac.jp, ††masunaga@is.ocha.ac.jp

Abstract　The Virtual World Database System (VWDB) is currently under development at Ochanomizu University. A preliminary design and implementation was done to support a multimodal query language in VWDB where users can issue query using mouse and voice. In this paper, we propose a strong VWDB multimodal query processing named a fusion query processing. By this processing users can issue queries with a condition relating to distance, direction , topology , and visibility as well as their combination.

Key words　VWDB，3-D database，Virtual reality，Multimodal，Query language, Signature

1. はじめに

我々は，仮想世界データベースシステム（Virtual World Database System:VWDB）の設計と実装を進めている．本システムは，バーチャルリアリティシステムとデータベースシステムを統合することによって，バーチャルリアリティシステムだけでは成しえなかった，仮想世界に対する問合せや更新処理時のトランザクション管理，一貫性の保持といったデータベース機能を十分に備えたシステムである．

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], [3] 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．さらに，仮想空間において，距離・方向・位相・可視性に関してそれぞれに問合せができる単純なマルチモーダル問合せ言語の開発が行われた [4]．さらに，VWDB システムに音声認識機能を搭載することで，ユーザは実世界に近い自然なインタラクションで問合せを発行することが可能となった．

しかしながら，複雑で本格的なマルチモーダル問合せ言語の

開発はまだなされていなかった．そこで本研究では，上記の問合せのうち２つの問合せを組み合わせることが可能な融合文法を定義し，それを用いた VWDB における複雑で本格的なマルチモーダル問合せ言語の開発について報告する．以下，2 章では VWDB システムの構成について，3 章では空間関係の問合せ MPQ とシグネチャによる表現，文法の生成について，4 章でマルチモーダル融合問合せ処理について説明し，5 章でプロトタイプシステムの実装について示し．6 章でまとめと今後の課題となっている．

2. 仮想世界データベースシステム：VWDB

VWDB のシステム構成やスキーマは問合せ処理と深くかかわるのでその概要を示す．

2.1 システム構成

VWDB はバックエンドのオブジェクト指向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，複数のユーザが仮想環境を共有できるようになっている [1], [2]．VR クライアントは，Java 言語用の仮想世界構築 API である Java3D を用いて実装し，バックエンド DB は商用のオブジェクト指向 DB である Object Store を利用し，Java 言語で実装している．さらに, 問合せインタフェースとして Advanced Media 社の AmiVoiceSDK の音声認識エンジンを用いてルール設計が進められている．

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．VWDB における仮想世界オブジェクトおよび仮想世界データベースのモデリングは基本的にオブジェクト指向設計手法に基づいて設計されている．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」とする．カテゴリはクラスの概念を継承する．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 カテゴリを継承し，さらに非空間的属性と形状属性をもつ．形状属性は継承せず，カテゴリごとに定義する．

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．

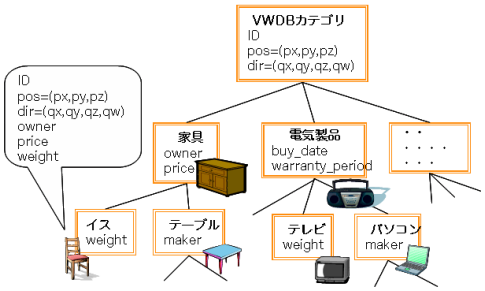

図 1　VWDB スキーマ

2.3 VWDB データベース空間

VWDB は，以上のようなスキーマをもとに構成された 3 次元のデータベース空間をサポートする．ユーザは仮想空間に没入している状態で，音声認識装置とマウスを用いたマルチモーダル操作によって仮想世界をウォークスルし，問合せを行う．図 2 は，VWDB システムのメインウィンドウを表している．

VWDB における問合せは，仮想世界オブジェクトの非空間的属性に対する問合せと，オブジェクト同士（または，オブジェクトとアバタ）の空間関係に関する問合せの 2 種類に分類することができる．

非空間的属性の問合せの例としては「10 万円以上のものは?」,「所有者がアキコさんの椅子は？」といった質問が挙げられる．

これらの問合せ述語は，前節で述べたオブジェクトのカテゴリをもとに記述することができる．

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（A1，A2，…，An）とする．A1，A2，…，An 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．

例えば,「イス」カテゴリのスキーマが「イス (owner, price, maker)」であるとき,「所有者がアキコさんのイスの値段は？」という問合せは,「イス (“ Akiko ”, ?, −)」という具合に記述できる．

3 次元空間ならではの問合せである空間関係に関する問合せには，空間関係の 3 大要素 [5] である距離・方向・位相に関する問合せ，およびアバタからの可視性に関する問合せが挙げられる．これらの空間関係の問合せ処理について，以下 3 章で説明する．

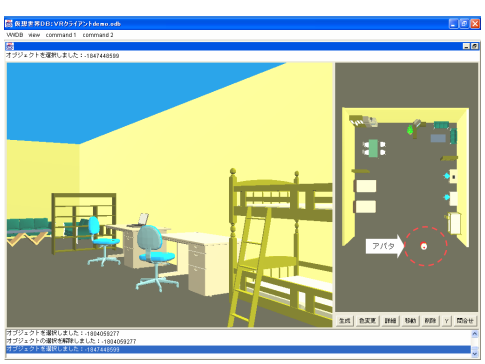

図 2　VWDB メインウィンドウ

3. 空間関係の問合せ

3.1 空間問合せ MPQ

空間関係に関する問合せ，例えば「このパソコンの右にあるものは？」といった問合せを処理する場合，空間関係を格納するリレーションが必要となる．

そこで，VWDB における空間関係の問合せとして，距離・方向・位相と，可視性を判断する 4 種類の空間関係を以下のように 3 項組または 4 項組の述語で表現し，これらに対応して空間関係のリレーションを用意する．

- 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(O_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 (=，<，≦，≧，>，≠)，dis：距離（数値），dir：(東・西・南・北・前・後・左・右・南東・南西・北東・北西・右前・左前・右後・左後)，top：位相（$meet$，$disjoint$，$inside$，$contains$），vis：$true/false$.

さらに，これらの述語に基づいた VWDB における意味のある基本的な空間問合せとして，図 3 にある 31 種の MPQ（meaningful primitive queries）が定義された [4].

3.2 シグネチャと文法

本節では，VWDB システムにおける空間関係の MPQ を基に，融合問合せの基礎的事項を述べる．まず，問合せ MPQ をもとに，データ型（data type）[6] の表現を用いて問合せを規定する．データ型は，sort，operator，constant から構成される．Sort により，VWDB における空間関係の問合せに用いられる要素を定義し，sort を用いて，問合せの関係性を表現したものが operator である．

MPQ	問合せ文の一例
(dis-1) Distance (O,O,=,?)	これとあれの距離は？
(dis-2) Distance (O,O,?,*dis*)	これとあれの距離はdisと比べてどうか？
(dis-3) Distance (O,?, *θ*,*dis*)	これとの距離がdis*θ*ものは？
(dis-4) Distance (?,O, *θ*,*dis*)	あれとの距離がdis*θ*ものは？
(dis-5) Distance (O,?,=,?)	これとすべてのものの距離は？
(dis-6) Distance (?,O,=,?)	あれとすべてのものの距離は？
(dis-7) Distance (O,?,?,*dis*)	これとすべてのものの距離はdisと比べてどうか？
(dis-8) Distance (?,O,?,*dis*)	あれとすべてのものの距離はdisと比べてどうか？
(dis-9) Distance (?,?, *θ*,*dis*)	距離関係がdis*θ*であるもの同士は？
(dis-10) Distance (?,?,=,?)	すべてのもの同士の距離関係は？
(dir-1) Direction(O,O,?)	あれはこれからみてどの方向にあるか？
(dir-2) Direction(O,?,*dir*)	これのdir (の方向)にあるものは？
(dir-3) Direction(?,O,*dir*)	あれがdir (の方向)にあるものは？
(dir-4) Direction(O,?,?)	これとすべてのものの方向は？
(dir-5) Direction(?,O,?)	あれとすべてのものの方向は？
(dir-6) Direction(?,?,*dir*)	dir の(方向)関係にあるもの同士は？
(dir-7) Direction(?,?,?)	すべてのもの同士の方向関係は？
(top-1) Topology(O,O,?)	あれはこれからみてどの位相にあるか？
(top-2) Topology(O,?,*top*)	これと top の(位相)にあるものは？
(top-3) Topology(?,O,*top*)	あれが top の(位相)にあるものは？
(top-4) Topology(O,?,?)	これとすべてのものの位相は？
(top-5) Topology(?,O,?)	あれとすべてのものの位相は？
(top-6) Topology(?,?,*top*)	top の(位相)関係にあるもの同士は？
(top-7) Topology(?,?,?)	すべてのもの同士の位相関係は？
(vis-1) Visibility(*Avt*,O,?)	あれは Avt からみえるか？
(vis-2) Visibility(*Avt*,?,*vis*)	Avtから vis なものは？
(vis-3) Visibility(?,O,*vis*)	あれが vis なアバタは？
(vis-4) Visibility(*Avt*,?,?)	Avt とすべてのものの可視関係は？
(vis-5) Visibility(?,O,?)	あれとすべてのアバタの可視関係は？
(vis-6) Visibility(?,?,*vis*)	vis の(可視)関係であるものとアバタのペアは？
(vis-7)Visibility(?,?,?)	すべてのものとアバタの可視関係は？

図 3　空間関係の問合せ（MPQ）

Operator σ，operand s_1〜s_n (sort)，return value s (sort) とすると operator は「σ ：$s_1 \times s_2 \times \ldots \times s_n \rightarrow s$」と表記できる．また，operand を持たない operator のことを constant といい「σ：$\rightarrow s$」と表す．これらのデータ型の表現を用いて，VWDB における問合せ MPQ から，距離シグネチャ・方向シグネチャ・位相シグネチャ・可視性シグネチャを規定した．以下に，シグネチャの例として，距離シグネチャを示す．

- **距離シグネチャ**
  - Sorts：

    $Object$, DIS,Θ, set($Object$), set($Object \times DIS$), set($Object \times \Theta$), set($Object \times Object$), set($Object \times Object \times DIS$)

  - Operators：

    o_i :$\rightarrow Object$ (1 ≦ i ≦ n)
    dis :$\rightarrow DIS$
    <:$\rightarrow \Theta$
    ≦:$\rightarrow \Theta$
    =:$\rightarrow \Theta$
    ≧$\rightarrow \Theta$
    >:$\rightarrow \Theta$
    ≠ :$\rightarrow \Theta$
    distance1:$Object \times Object \rightarrow DIS$
    distance2:$Object \times Object \times DIS \rightarrow \Theta$
    distance3:$Object \times \Theta \times DIS \rightarrow$set($Object$)
    distance4:$Object \times \Theta \times DIS \rightarrow$set($Object$)
    distance5:$Object \rightarrow$set($Object \times DIS$)
    distance6:$Object \rightarrow$set($Object \times DIS$)
    distance7:$Object \times DIS \rightarrow$set($Object \times \Theta$)
    distance8:$Object \times DIS \rightarrow$set($Object \times \Theta$)
    distance9:$\Theta \times DIS \rightarrow$set($Object \times Object$)
    distance10:$\rightarrow$set($Object \times Object \times DIS$)

ただし，o_i は仮想空間内にある有限個のオブジェクトを意味する（i ≦ n）．dis はオブジェクト同士の距離の小数点第三位を四捨五入した離散値（$\in DIS$）とする．

ここで，operator である distance 1，distance 5，distance 6，distance 10 の operand に，MPQ で用いられていた Θ がないことに気づく．これは distance1，distance 5，distance 6，distance 10 の問合せは，Θ が'='のときのみ意味のある問合せとして定義されているからである．

Sort S を用いた operator のセットのことを S-sorted シグネチャという．距離に関しては，S=$Object$，DIS，Θ，set($Object$)，set($Object \times DIS$)，set($Object \times \Theta$)，set($Object \times Object$)，set($Object \times Object \times DIS$) となる．

次に，これらのシグネチャから文法を定義する．Operator で

構成される表現を term という．Term は次のように定義されている．

(a) constant $\sigma :\rightarrow s$ に対して，表現 "σ "は term である．

(b) operator $\sigma : s_1 \times s_2 \times \ \times s_n \rightarrow s$ と sort s_i の term t_i ($1 \leqq i \leqq$ n) に対して，表現 "$\sigma(t_1 \times t_2 \times \times t_n)$ " は term である．

全ての term を生成することができる文法をシグネチャから構成することができる．つまり，各 sort は非終端記号に置き換えられる．

各 operator「σ: $s_1 \times s_2 \times \ \times s_n \rightarrow s$（n>0）」は生成則「$s \rightarrow \sigma(s_1, s_2,, s_n)$」に置き換えられ，各 constant「$\sigma$: $\rightarrow s$」は生成則「$s \rightarrow \sigma$」と置き換えられる．したがって，距離・方向・位相・可視性に関して，それぞれ文法が定義できる．以下に例として距離文法を示す．

- **距離文法**

$Object \rightarrow o_i$ ($1 \leqq i \leqq$ n)
$DIS \rightarrow dis$
$\Theta \rightarrow <$
$\Theta \rightarrow \leqq$
$\Theta \rightarrow =$
$\Theta \rightarrow \geqq$
$\Theta \rightarrow >$
$\Theta \rightarrow \neq$
$DIS \rightarrow$distance1($Object, Object$)
$\Theta \rightarrow$distance2($Object, Object, DIS$)
set($Object$) $\rightarrow$distance3($Object, \Theta, DIS$)
set($Object$) $\rightarrow$distance4($Object, \Theta, DIS$)
set($Object \times DIS$) $\rightarrow$distance5($Object$)
set($Object \times DIS$) $\rightarrow$distance6($Object$)
set($Object \times \Theta$) $\rightarrow$distance7($Object, DIS$)
set($Object \times \Theta$) $\rightarrow$distance8($Object, DIS$)
set($Object \times Object$) $\rightarrow$distance9(Θ, DIS)
set($Object \times Object \times DIS$) $\rightarrow$distance10

3.3 問合せ構文木

シグネチャから導出された文法によって，各 sort は非終端記号に置き換えられるため，VWDB における問合せは，例えば図 4，5，6 に示すような構文木を形成することになる．図 4 は，「これ（o_1）の南の方向にあるものは？」という問合せに，図 5 は,「o_1 と o_2 の距離は dis と比べてどうか？」という問合せに対応する．さらに，図 5 の DIS をさらに展開したものが図 6 で,「o_1 と o_2 の距離は o_3 と o_4 の距離と比べてどうか？」という問合せに対応する．このような構文木が，距離・方向・位相・可視性に関する問合せについてそれぞれ形成される．

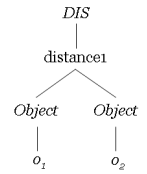

図 4 構 文 木 1

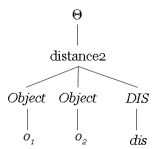

図 5 構 文 木 2

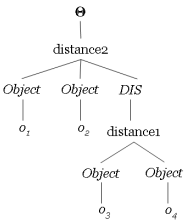

図 6 構 文 木 3

4. マルチモーダル融合問合せ

本章では，距離・方向・位相・可視性に関する個々の問合せに加えて，それらを融合させた問合せ（融合問合せと呼ぶ）を処理するための融合文法を定め，それを用いたマルチモーダル融合問合せ処理法について論じる．

4.1 融合文法の導入

図 6 の問合せ構文木は，$\Theta \rightarrow$distance2($Object$,$Object$,DIS) と $DIS \rightarrow$distance1($Object, Object$) を Sort DIS を介してつなぎ合わせた問合せといえる．このように，つなぎ合わせることが可能な構文木同士というのは，生成則でみると右辺の Sort と左辺の Sort が等しい生成則同士であることが分かる．

距離・方向・位相・可視性に関する構文木同士も同様に，共通する Sort を介在することで，問合せを融合できると考えられる．距離・方向・位相・可視性シグネチャに共通する sort として $Object$ と set($Object$) が挙げられる．ここで，生成則の右辺に operator の変数として $Object$ が出現するが，$Object$ を左辺に持つ生成則は $Object \rightarrow o_i$ という明らかな生成則以外にはないことに気付く．しかし，set($Object$) を左辺に持つ生成則は，構文木を評価して得られる一般には複数のオブジェクトの集合とみなせば，それを更に適用できると考えられる．例えば，生成則 $DIS \rightarrow$ distance1($Object, Object$) の右辺に現れた $Object$ を介在して，生成則 set($Object$) $\rightarrow$ visibility3($Avt \times Bool$) をさらに適用していく．これが，融合の意味である．

実際，set($Object$) を左辺に持つ生成則を評価して得られたオブジェクト集合（ここでは，"これら"と書く）とあるオブジェクト（ここでは，"あれ"と書く）について「これらとあれの距離は?」という質問は意味がある．従って，$Object$ と set($Object$) を結合する文法規則を持つことで，距離方向位相可視性文法を融合する「融合文法」を導入することができる．

- **融合文法**

以下の 2 種類の生成則からなる文法と定義する．

(1) 距離・方向・位相・可視性文法の全ての生成則
(2) 融合のための生成則
$Object \rightarrow$fusion(set($Object$))

図 7 に，融合問合せ文法を用いた融合問合せの構文木の一例を示す．この構文木は以下の二つの文法を融合した問合せである．

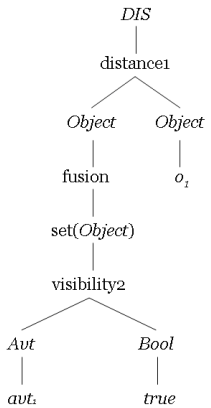

図 7　融合問合せ構文木 1

- 可視性文法：set($Object$) →visibility2($Avt, Bool$)
- 距離文法：DIS →distance1($Object, Object$)

この構文木は「avt_1 から見えるものと o_1 の距離は？」という問合せを意味する．このように，融合文法を用いることにより，空間関係の融合問合せが表現可能となる．

4.2　融合問合せの組合せパターン

本節では，どのような融合問合せが実際に融合文法を用いて生成されるか説明する．融合文法によって，右辺に $Object$ をもつ生成則と左辺に set($Object$) をもつ生成則を融合することができる．

以下にあるように，右辺に $Object$ をもつ生成則は 21 種類，左辺に set($Object$) をもつ生成即は 7 種類存在する．

- 右辺に $Object$ をもつ生成則
  - 距離に関する生成則

DIS →distance1($Object, Object$)
Θ →distance2($Object, Object, DIS$)
set($Object$) →distance3($Object, \Theta, DIS$)
set($Object$) →distance4($Object, \Theta, DIS$)
set($Object$, DIS) →distance5($Object$)
set($Object$, DIS) →distance6($Object$)
set($Object$,Θ) →distance7($Object, DIS$)
set($Object$,Θ) →distance8($Object, DIS$)

  - 方向に関する生成則

DIR → direction1($Object$, $Object$)
set($Object$) →direction2($Object$, DIR)
set($Object$) →direction3($Object$, DIR)
set($Object$, DIR) →direction4($Object$)
set($Object$, DIR) →direction5($Object$)

  - 位相に関する生成則

TOP → topology1($Object$, $Object$)
set($Object$) →topology2($Object$, TOP)
set($Object$) →topology3($Object$, TOP)
set($Object$, TOP) →topology4($Object$)
set($Object$, TOP) →topology5($Object$)

  - 可視性に関する生成則

$Bool$ →visibility1(Avt, $Bool$)
set(Avt) →visibility2($Object$, $Bool$)
set(Avt, $Bool$) →visibility4($Object$)

- 左辺に set($Object$) をもつ生成則
  - 距離に関する生成則

set($Object$) →distance3($Object, \Theta, DIS$)
set($Object$) →distance4($Object, \Theta, DIS$)

  - 方向に関する生成則

set($Object$) →direction2($Object$, DIR)
set($Object$) →direction3($Object$, DIR)

  - 位相に関する生成則

set($Object$) →topology2($Object$, TOP)
set($Object$) →topology3($Object$, TOP)

  - 可視性に関する生成則

set($Object$) →visibility3(Avt, $Bool$)

距離・方向・位相・可視性に関する問合せのうちの 2 つを組み合わせた融合問合せは，以上の生成則から 147 種類生成されることとなる．さらに，DIS →distance1($Object, Object$) のように，右辺に $Object$ をもつ生成則で，ひとつの生成則中に $Object$ が二つ存在するものもある．この場合，どちらの $Object$ にも，融合文法は適用可能であるため，融合文法を用いた問合せは，175 種類生成されることとなる．

しかし，これらの融合問合せを意味的に捉えると，同一の問合せを示しているものが存在することに気付く．たとえば，図 8 と図 9 の融合問合せ構文木は共に「これとの距離が $dis\theta$ のものと全てのものの距離は？」という問合せを意味している．

しかし，各問合せ構文木を構成している融合文法以外の生成則は順に set($Object$, DIS) →distance5($Object$) と set($Object$)

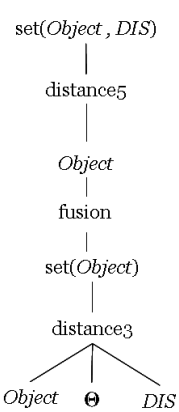

図 8　融合問合せ構文木 2

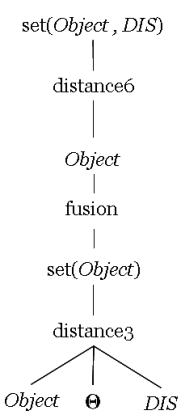

図 9　融合問合せ構文木 3

→distance3($Object$, Θ, DIS)，および set($Object$, DIS) →distance6($Object$) と set($Object$) →distance3($Object$, Θ, DIS) である．

MPQ を規定した際の基点オブジェクトと対象オブジェクトの違いが，生成則 set($Object$, DIS) →distance5($Object$) と生成則 set($Object$, DIS) →distance6($Object$) の違いである一方で，これらの生成則の右辺の $Object$ に融合文法を適用し，融合問合せを構築した場合，結果として同じ問合せを意味することになる．

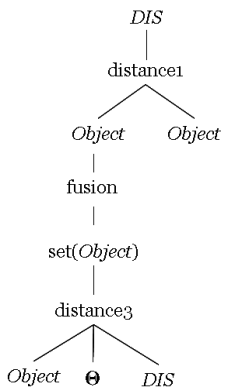

図 10　融合問合せ構文木 4

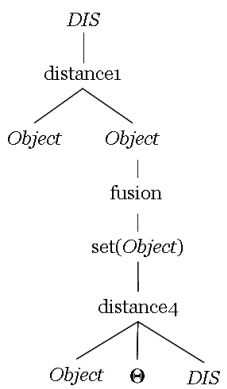

図 11　融合問合せ構文木 5

さらに，図 10 と図 11 の融合問合せ構文木をみると，それぞれ問合せ文は「これとの距離が $dis\theta$ のものとあれの距離は?」，「あれとの距離が disθ のものとこれの距離は?」という具合に，「これ」,「あれ」という違いはあるものの，指し示すオブジェクトが同じものであった場合，いずれの問合せも同じ問合せ結果を返す．このような問合せを考慮した結果，175 種の融合問合せのうち意味的に独立している問合せは 103 種類となる．

4.3　マルチモーダル融合問合せ処理

仮想世界内で，ユーザは音声とマウスによる入力で問合せを入力する．本節では，融合問合せ処理のシステム構成を示し，ユーザが発声した融合問合せが，前章で示された 103 種の問合せのうちのどの問合せであるかを判別する音声認識ルールについて論じる．

4.4　システム概要

VWDB では，音声認識エンジンとマウスによるポインティングを用いたマルチモーダル問合せ処理により，ユーザは SQL のような特別な問合せ言語を使用せずに，より実世界に近い形で問合せを行う．ユーザは，マウスの操作によって仮想世界を自由に動き回り，オブジェクトの生成・削除・プロパティの参照・色変更などを行うことが可能である．

図 12 に，マルチモーダル融合問合せ処理システム構成図を示す．マルチモーダル問合せ処理では，ユーザは音声で問合せを発声すると共に，問合せの対象オブジェクトをマウスでクリックし，オブジェクトのコードを時間順序に合わせてピッキングデータとして取得する．

VWDB 空間では常に（問合せとは関係なく）ユーザのマウスイベントを監視している．マウスで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）を時間順に格納しておく．音声認識されたキーワード中に「これ」,「あれ」といった，オブジェクトを示す語がが含まれていた場合は，そのキーワードの個数によって次のように処理する．

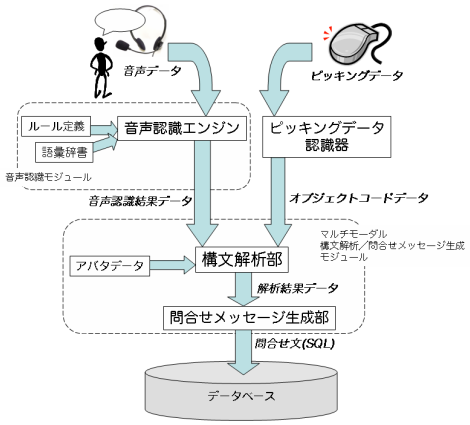

図 12　マルチモーダル問合せ処理システム構成図

- 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．

- 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．

このようにして取得されたピッキングデータと，音声認識エンジンによって認識された問合せ文を統合することにより，ユーザがどのオブジェクトに対し，どのような問合せを発声したのか解析を行い，SQL に変換を施しデータベースへ問合せが成される．

音声認識処理法には，ディクテーション認識による複雑な自然言語処理と，事前に語彙辞書に登録したセンテンスのみを認識するルール認識処理がある．VWDB における問合せ処理では,，使用されるキーワードがある程度限定されるため，ルール認識方式の方が的確かつ効率的に処理できると考えられる．

4.5　音声認識モジュール

ルール認識により，ユーザの発声した問合せを認識し，その中のキーワード集合をあらかじめ用意されている問合せライブラリを参照することで，問合せの認識を行う．各融合問合せに固有な非終端記号の集合を R とし，R を問合せライブラリにあらかじめ格納する．各 R は互いに同一とならないようにルールを設計できるので，ユーザからの音声認識結果と R のパターンマッチで各融合問合せへの対応付けが可能となる．

図 13 に，距離と距離の融合問合せの例と，それに対応する非終端記号の集合 R を示す．例えば,「これとの距離が 5m よりも小さいものと，あれの距離は？」という融合問合せを発声した場合,「これ（obj1），距離（DIS），5（dis），より小さい

融合問合せ文	非終端記号の集合R
A：これとの距離がdisθのものとあれの距離は？	{obj1,DIS,dis,θ,obj2,DIS}
B：これとの距離がdisθのものとあれの距離はdisと比べてどうか？	{obj1,DIS,dis,θ,obj2,DIS,dis}
C：これとの距離がdisθのものとの距離がdisθのものは？	{obj1,DIS,dis,θ,DIS,dis,θ}
D：これとの距離がdisθのものとすべてのものの距離は？	{obj1,DIS,dis,θ, DIS}

図 13　問合せ文とそれに対応する非終端記号の集合 R

(θ)，あれ（obj2），距離（DIS）」というキーワードを認識し，この集合 {obj1，DIS，dis，θ，obj2，DIS} と，問合せライブラリに格納してある R のパターンマッチを行い，図 13 の A の問合せであると判断される．さらに,「これ」に対応するオブジェクトのコードと「あれ」に対応するオブジェクトのコードを順に取得し，認識した問合せと関連づけて SQL に変換する．

4.6　融合問合せの発声

マルチモーダル融合問合せ処理では，各問合せ文に固有な集合 R に基づいて，音声から問合せ文の認識を行うため，ユーザが発声した言葉の順序が大きな意味をもつ．そのため，発声のルールとして，複数の構文木から問合せが構成されている場合，下層の問合せ構文木から問合せを発生し，問合せを行うことと定める．例えば，図??の問合せ構文木の場合，$Object$ →visibility2($Avt, Bool$) という下層の構文木の問合せの次に DIS →distance1($Object, Object$) という上層の構文木が発声され,「avt_1 から見えるものと o_1 の距離は？」という問合せ文となる．

4.7　音声認識ルール

空間関係の問合せを特徴付ける単語をあらかじめ定義し，問合せに対応したルールを設計する．ルールの記述は拡張 BNF 構文（ABNF）の一種である JSpeech Grammar Format（JAGF）に従っている．以下に，距離と距離の融合問合せ「これとの距離が $dis\theta$ のものとあれの距離は？」の音声認識ルール記述例を示す．

- **融合問合せルール例**

  問合せ：これとの距離が $dis\theta$ のものとあれの距離は？

```
public<fusion-1-A>=[<start>] <obj1> <DIS>
 <dis> <theta> <and1> <obj2> <DIS> <stop>;
 ----------------------------------------
  <start> = (この中で|これらの中で|
             これらのうちで);
   <obj1> = (これ|この人| 私|自分)<tofrom>;
 <tofrom> = [から][まで][との][と][は][の]
             [みて][を][基準][と][して][が];
    <DIS> = 距離 [関係][が][は];
    <dis> = <NULL>{num}(<int>|<float>)
             {num}[メートル];
  <theta> = [より](小さい|大きい|以上|以下|
             以内|の|のところ|である|
             の関係にある|同士である|同士の|
             でない|離れている|近い|遠い|
             近く|遠く)
             [離れている][遠く];
  <and1> = [ものと];
  <obj2> = (あれ|あの人)<tofrom>;
  <stop> = [の][関係][に][である][みえる]
             [もの][オブジェクト][の][ペア]
             [同士][は][どれ][どの]
             [どのような][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][ろ][て]
             [で][あり][くれ][ます][か]
             [ちょうだい][下さい][な];
```

<> はルール名を表し，| は”または”を表す．[ ] は，その発音があってもなくてもよいものを表す．この融合問合せの非終端記号の集合は {obj1，DIS，dis，θ，obj2，DIS} であるが，実際に音声認識を構築する際は，さらに助詞などを考慮して，より自然な日本語で問合せが行えるような記述を行う．

このようなルール設計により,「これとの距離が 8 メートルより遠いものとあれの距離を示せ」という具合に発声した場合でも,「これ」,「距離」,「8」,「より遠い」,「あれ」,「距離」というキーワードを認識することで，<fusion-1-A> の問合せであると認識することが可能となる．

5.　プロトタイプシステム

ユーザは VWDB を起動させ，マウス操作によって，仮想世界をウォークスルーする．問合せを行う時は，VWDB 画面上の，問合せボタンをクリックすることで，問合せ画面がたち上がり，問合せを発声すると共に，対象となるオブジェクトのクリックを行う．仮想世界オブジェクト全体が対象となるような問合せの場合は，クリックを行わずに，問合せのみを発声する．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（オブジェクトの色を変更し，数秒間点滅），オブジェクトに付加した 3 D テキストによる表示（距離や方向などを示す結果を伴う場合），および問合せウィンドウにテキスト表示で行う．

しかし，融合問合せでは，問合せ結果を一度に描写することが困難な問合せが多く見られる．例えば,「これとの距離が 1m 以内のものの東にあるものは？」という融合問合せでは，まず「これとの距離が 1m 以内のもの」という問合せを評価し，その結果，該当するオブジェクトが k 個（obj_1〜obj_k）存在するとする．ただし，n は仮想世界中のオブジェクト数 (0≦ k ≦n) とする．次に，それらの該当オブジェクト obj_1〜obj_k 各々に対し,「これの東にあるものは？」という問合せを行う．

「これ（obj_1）の東にあるものは？」の問合せに該当するオブジェクトを obj_{11}〜obj_{1n},「これ（obj_2）の東にあるもの

は？」の問合せに該当するオブジェクトを obj_{21}〜obj_{2m}，「これ（obj_k）の東にあるものは？」の問合せに該当するオブジェクトを obj_{k1}〜obj_{kl} とする（$0 \leq m, l \leq n$）と，融合問合せ結果として該当するオブジェクトはこれら全てを指すこととなり，同一画面上で表示するには限界がある．そこで，融合問合せの結果表示では，obj_1〜obj_k をリストでまず示し，リストから obj_1 を選択すると，VWDB 画面に結果として obj_{11}〜obj_{1n} が強調表示され，リストから obj_k を選択すると，obj_{k1}〜obj_{kl} が強調表示されるようにした．融合問合せの結果表示の例を図 14，15，16 に示す．これらは，融合問合せ「これとの距離が 3m 以内のものとあれの距離は？」の結果画面の概観である．

実際にオブジェクトを 30 ほど配置した VWDB 空間で，アバタは現在，空間の中心付近にいる．ユーザは融合問合せ文を発声するとき,「これ」と発声すると同時に画面左にあるテーブルをクリックし,「あれ」と発声すると同時に手前にあるベッドをクリックする．まず，テーブルの 3 m以内に存在しているオブジェクトは 11 個存在するという結果が画面右上の問合せウィンドウに表示され，それらのオブジェクトコードがリストで表示される（図 14）．

次に，リストから一つコードを選択すると，その選択されたオブジェクトとベッドが強調表示され，続けて距離が表示される．図 15 では本棚を選択した場合の結果，図 16 ではテレビを選択した場合の結果をそれぞれ強調表示している．図 16 のように，ユーザの視界から見えないオブジェクトが選択された場合においても，画面右側の鳥瞰図によりオブジェクトの位置を把握することができる．

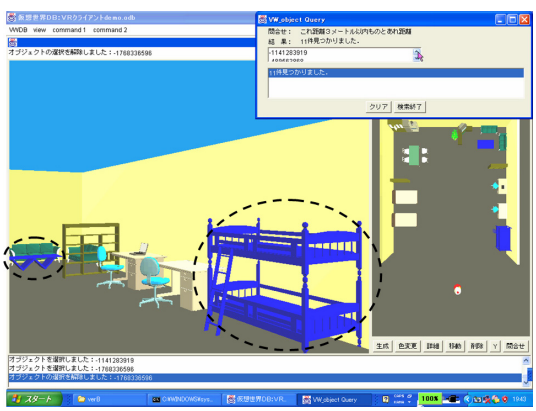

図 14　融合問合せ例：オブジェクトを二つ選択

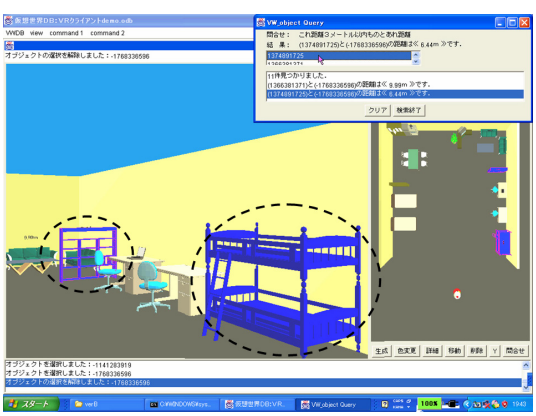

図 15　融合問合せ結果例：本棚とベッドの距離を表示

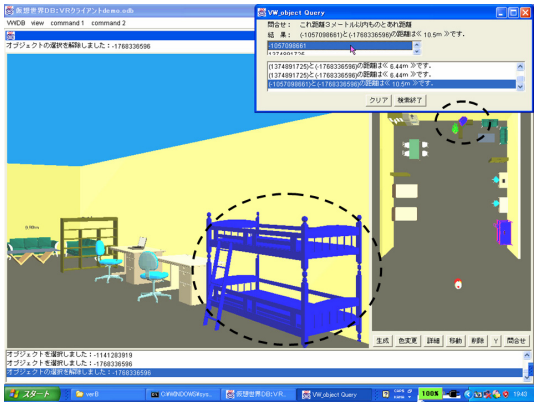

図 16　融合問合せ結果例：テレビとベッドの距離を表示

以上のようにして，VWDBにおける融合問合せ処理が音声とマウス操作によって行われることが示された．

6.　まとめと今後の課題

本稿では VWDB におけるマルチモーダル問合せシステムの拡張であるマルチモーダル融合問合せシステムの構築を行った．距離・方向・位相・可視性のうち二つを融合することができる融合文法を用いた融合問合せの種類を示し，それらを音声で認識するためのルールを設計した．また，問合せ結果の表示法について論じ，音声認識エンジンとマウスによる操作で問合せを行い，その結果を示した．今後の課題として，二つ以上の問合せなど，より高度な問合せへの対応，サイバーグローブによるジェスチャの認識を導入で，マウスに依存しないポインティング操作の実現などが挙げられる．

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] 渡辺知恵美，増永良文，仮想世界データベースシステム VWDB2 における仮想世界同期法，情報処理学会誌：データベース (TOD)，TOD18，2003.

[4] 矢野ナホコ，横川明子，増永良文，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，日本データベース学会 Letters，Vol.4，No.1，pp. 5-8，2005.

[5] M. Egenhofer, Spatial Relation: Models, Interfaces, and their Future Application, Proceedings of Advanced Database Symposium '96, SIGDBS, IPSJ, 31p., 1996.

[6] J. Craig Cleaveland: An Introduction to Data Types, AT & T Bell Laboratories, 239p., 1986.