個別化されたファジィデータ管理のための グラフ指向スキーマ設計†

藤代 一成*¹ 白井 靖人*² 池辺 八洲彦*¹ 國井 利泰*³

大量のあいまいな情報の効率的な一括処理・管理を目的として、ファジィ集合論に基づくデータベースシステム—ファジィデータベースシステム (FDBS) に関する研究・開発が活発に行われている。客観的な共有データに加え、個別化されたファジィデータも統合・管理し、それらに基づく主観的な問合せの処理を可能にすることは、FDBSに関する中心的課題の一つであると考えられる。

本論文では、可能性分布を値として許す個別化されたファジィ属性データを管理する方式を提案する。そのために、まずグラフ指向的な疑似意味データモデルである拡張グラフデータモデル (EGDM) を用いて、上述のファジィ属性データを効果的に表現するスキーマ設計法を開発する。次に、このスキーマ表現に基づいて、複数ユーザのファジィ属性データを統合的に管理し、ユーザごとに個別化された実現値を提供するようなユーザビュー＝主観的ファジィ実現値ビューを実現するための、FDBSの基本アーキテクチャと問合せ処理手続きを提案する。さらに、年齢に関する簡単な例データベースを利用し、個別化された可能性分布の違いが、同一の対象世界に対する同一の問合せ評価に対し、異なる結果を導くことを具体的に示す。

キーワード：ファジィデータベースシステム、グラフ指向データモデル、可能性分布、個別化、主観的ファジィ実現値ビュー、主観的ファジィデータストレージ

1. はじめに

1.1 背景

大量のあいまいな情報の効率的な一括管理・処理を目的として、ファジィ集合論 [1] に基づくデータベースシステム—ファジィ・データベースシステム (FDBS) に関する研究・開発が活発に行われている [2] [3]。情報にあいまいさを持ち込む要因の一つに考えられるものに、個人の主観がある。そこで、対象の客観的な共有データに加え、個々の主観によって個別化されたデータの集まり

も統合的に管理し、それらに基づく主観的な問合せの処理を可能にすることは、FDBSに関する中心的課題の一つであると考えられる。

これまでに、さまざまな種類のファジィデータを扱うための拡張関係モデル、あるいは言語・システムが提案されてきている [4] [5] [6] [7] [8] [9] [10]。その中でも、データの個別化 (individualization) を扱う代表的な試みとして、Zemankova-Leechらのファジィ・関係データベース (FRDB) [11] [12] をあげることができる。彼女らは、FRDBを

- a) 値データベース (VDB)
- b) 説明データベース (EDB)
- c) 変換規則

の3つの独立したパートに分割し、ユーザの主観による局所的なEDBの変更に基いて、FRDB全体を個別化する考え方を提案した。EDBには、VDBに属する関係スキーマの属性値として許さ

† The Design of a Graph-Oriented Schema for the Management of Individualized Fuzzy Data
 Issei FUJISHIRO, Yasuto SHIRAI, Yasuhiko IKEBE and
 Toshiyasu L. KUNII

*¹ 筑波大学 電子・情報工学系
 Institute of Information Sciences and Electronics, University of Tsukuba

*² 静岡大学 教育学部
 Faculty of Education, Shizuoka University

*³ 東京大学 理学部
 Faculty of Science, The University of Tokyo

れる可能性分布ラベルの定義を始め、意味関係、類似関係、近接関係、一般ファジィ関係等の実現値が格納される。しかし彼女らは、個別のEDB実現値の集まりを統一的に管理する方式について、実装の観点から明確に記述していない。さらに、VDBの実現値や変換規則にも、同様の個人差が許されるべきであると考えられる。

1.2 目的と構成

そこで本論文では、属性値に可能性分布を許すようなクラス [8] の属性データに限定して、個別化されたファジィデータを統合的に管理する方式を提案する。そのために、まずグラフ指向的な疑似意味データモデルである拡張グラフデータモデル (EGDM) の定義をさらに拡張し、上述のクラス of ファジィ属性データを重み付き有向グラフによって効果的に表現するスキーマ設計法を開発する。次に、このスキーマ表現に基づき、複数ユーザのファジィ属性データを統一的に管理し、ユーザごとに個別化された実現値を提供するようなユーザビュー＝主観的ファジィ実現値ビューを提供するための、FDBSの基本アーキテクチャと問合せ処理手続きを提案する。

まず次節では、本論文で提案する主観的実現値ビューの基本的な役割と概念を関係モデルを用いて明らかにする。次に、3節でEGDMによるファジィ属性データのスキーマ表現を示した後、4節でEGDM上に実現される主観的ファジィ実現値ビューの管理方式を提案する。なお、3節と4節では、年齢に関する簡単な例データベースを用いて、個別化された可能性分布の違いが、同一の対象世界に対する同一の問合せの評価に対して、異なる結果を導くことを具体的に示すことにする。最後に5節で、本論文の総括と今後の研究方針を述べる。

なお本論文は、文献 [13] の研究会における口頭発表の内容を発展させたものである。

2. 主観的実現値ビューの基本的な概念

本節の目的は、従来からファジィデータモデリ

ングの基礎モデルとして頻繁に利用されてきた関係モデル [14] を用いて、次節以降重要な意味を持つ主観的実現値ビューの役割を、従来のビューのそれと明確に区別し、併せてその基本的な概念を具体的に示すことである。

2.1 従来型のビュー

通常、データベースシステムにおけるビューは、概念スキーマの単なる部分集合に相当する場合 (サブスキーマとも呼ばれる) を除けば、ユーザの利用環境に応じて、概念スキーマに含まれるいくつかの基本スキーマから、自明ではない規則によって、仮想的な型として導出されるものと定義される。

まず、関係データベースにおける典型的なビューの例を図1に示す。同図は標準的な略記法によって記述されており、関係スキーマ (relation schema) 名に続く () 中に、構成する属性名リストを記している。下線を付した属性名は、関係スキーマの主キーであることを意味する。なお、以下の図においても同様の略記法を用いることにする。

R_1 = 従業員 (氏名, 入社年度, 部署名)

R_2 = 部署 (部署名, 所在地, 管理者名)

R_3 = 管理一覧 (氏名, 管理者名)

R_4 = 従業員 89 (氏名, 部署名)

図1 関係ビューの例

図1では、 R_1 と R_2 が基本関係スキーマである。ビュー R_3 は、 R_1 と R_2 から関係代数式：

$$R_3 = \pi_{\text{氏名, 管理者名}} (R_1 \bowtie R_2)$$

によって、その実現値に依らず導出可能である。ここで、 π は射影演算子、 $\bowtie$ は自然結合演算子である。 R_1 は人事部門、 R_2 は総務部門にとってそれぞれ有用であるが、管理者が部下の氏名を掌握するという応用に限れば、ビュー R_3 は過不足ない情報を表現していると考えられる。

また、属性値に関して一定の制約を設けることによって、単一の関係の実現値を絞る役割を果たすビューも存在する。例えば図1の「1989年度入

社の従業員」を求めるビュー R_4 は、以下のような導出規則に従う：

$$R_4 = \pi_{\text{氏名, 部署名}}(\sigma_{\text{入社年度}=1989}(R_1))$$

ここで、 σ は選択演算子を表す。

さらに、複雑なデータ型を持つ概念(意味)データモデルにおける意味的相対性(semantic relativism) [15] も、応用ごとに、同一の対象を異なる型として見なすビュー機能の抽象化に他ならない。

2.2 主観的実現値ビュー

このように、応用上必要となる仮想的な型を導出することを目的とする従来型のビューとは異なり、対象世界の構造に対する見方は一定であっても、それを記述する属性に対して、個人の主観がもたらす実現値の多様性を許すような値指向のビューも考えられる。

例として、図2に示すような科目の難易度を扱う関係スキーマを考える：

S_1 = 学生 (氏名, 学年)

S_2 = 科目 (科目名, 単位数)

S_3 = 難易度 (科目名, 氏名, グレード)

図2 難易度に関する関係スキーマ

科目ごとに、ある基準に沿って客観的に決められた単一の難易度を表現するだけならば、 S_3 は不要であり、代わりに S_2 を拡張して、

S_4 = 科目 (科目名, 単位数, グレード)

を用意すれば十分である。しかしここでは、 S_3 によって学生ごとに異なる主観的な難易度グレードが一括保持されているので、 S_1 に含まれる個々の学生は、自分の氏名を検索キーの一部に使い、 S_3 から、自らの主観的な難易度データを引き出すことができる。実際、氏名 x の関数として、関係スキーマ $S_3[x]$ を：

$$S_3[x] = \pi_{S_4}(\sigma_{\text{氏名}=x}(S_2 \bowtie S_3))$$

によって定義すれば、 x の値によって相異なる難

易度をもつ S_4 と同型の関係を複数得ることができる。この $S_3[x]$ が、本論文で提案する主観的実現値(subjective occurrence)ビューの例に他ならない。 $S_3[x]$ を、前項で示した従来型のビューと区別する理由は、 S_4 と全く同型でありながら、実現値が異なる $S_3[x]$ の関係を、相異なるビューとして扱うからである。ここで、主観的実現値ビューを構成するために、主観的なデータを一括保持している関係スキーマ S_3 を、主観的データストレージ(subjective data storage)と呼ぶことにする。

本論文ではこれ以降、断りのない限り、ビューはこの主観的実現値ビューを意味するものとする。

3. EGDMによるファジィデータの表現

図2からも明らかのように、関係モデルは、その定義の簡潔さの反面、主観的実現値ビューを構成するための各関係スキーマの役割が明確に区別できないという欠点をもつ。これは、個別化されたデータを統合的に扱うデータベーススキーマを論理設計していく上で、大きな支障になると考えられる。そこで本節では、関係モデルの代わりに、文献[16][17]で著者らがこれまで開発してきたデータモデルEGDMの定義の一部をさらに拡張し、それを用いてビュー構成の基本となるファジィ属性データのモデリングを行うことにする。なお、本節3項で述べるモデリング方式は、文献[18][19]の方式を改良したものである。

3.1 可能性分布属性の関係表現

以下では、ファジィ属性として可能性分布を許すようなクラス[8]のファジィデータ表現を試みる。そこでまず、簡単に可能性分布[20][3]について触れておくことにする。

対象世界(universe of discourse) U に含まれる任意の実体要素 x の単純属性 A の値 $A(x)$ が、あいまいな単語 P で述べられている命題：

$$[A(x) \text{ は } P \text{ である}]$$

が与えられたとする。このとき、 x の属性 A に関する可能性分布は、

$$\Pi_{A(x)} = \begin{cases} P & (3.1) \\ \sum_{i=1}^n \pi_A(u_i) / A(u_i) & (3.2) \end{cases}$$

で定義される。ここで、 $\pi_A(u_i)$ ($i=1, \dots, n$) は、要素 A に関する実体要素 u_i の可能性分布関数と呼ばれ、各実体要素 u_i の属性値の可能性

$$\pi : A(U) \rightarrow [0, 1]$$

を決めるものである。通常 (3.2) 式には、可能性 = 0 の属性値は含めない。

図 3(a) に年齢に関する関係データベースの例を示す。ここでは、各人の年齢がファジィ属性になっている。同図からも明らかなように、可能性分布属性には 2 種類の表現方法がある。一つは、"young" のような式 (3.1) に対応する方法である。これは、一つの可能性分布を代表する一種のラベルであり、この他にも、式 (3.2) に対応する可能性分布を表す具体的な定義が必要になる。もう一つは、式 (3.2) に対応する可能性分布そのものが直接、属性値になっている場合である。この場合のタプルは、多価の複合属性を含むので、このままでは第一正規形の条件を満足しないことになる。

また同図には、"unknown" を意味する空値 (null values) [21] が現れている。こうした空値に対しても、関係モデルでは特別の取扱いが必要である [8]。

3.2 GDMとEGDM

関係モデルの定義の簡潔さと数学的取扱いの容易さを活かしながら、データインテンシブな応用を扱うための意味的記述能力を改善し、より高度な概念モデリングを可能とするような後続のデータモデルが数多く提案されてきている。

その一つに、1983 年に國井によって発表されたグラフ・データモデル (GDM) がある [22]。GDM では、実体型を表現する関係スキーマをレコード型と呼ぶ。そして、二つの (異なっている必要はない) レコード型間の多対多関連型を、双方向のリンク型によって表現することにより、関係モデルの

PERSON

NAME	SEX	AGE
James	male	young
Linda	female	0.5/18+0.9/19+0.8/20
Joseph	male	unknown

YOUNG

AGE	GRADE
17	0.7
18	1.0
19	0.9

ADULT

AGE	GRADE
20	1.0
40	1.0

MIDDLE

AGE	GRADE
40	1.0

(a) 関係

(b) EGDMスキーマ

図 3 可能性分布によるファジィ属性データの関係モデルおよびEGDMによる表現

結合演算と比べて、複数のレコード型に関わるデータアクセスの処理効率を向上させるばかりでなく、実体と関連とを明確に区分したデータベースモデリングの環境をユーザに提供できるようなモデルである。GDMスキーマは、レコード型をノード、リンク型をアークとする有向グラフによって表現できるため、グラフ指向、あるいはリンク指向、パス指向と呼ばれている [23]。GDMは、単なる概念モデルにとどまらず、それに基づく商用DBMS、G-BASEも既に開発されている [24] [25]。

1987 年に著者らは、このGDMに対して、リンク型自身がもつ属性、すなわちリンクフィールドを許すような拡張を施し、E (xtended) GDMを定義し、その実装方式を提案した [16]。結果として、

EGDMは実体-関連(Entity-Relationship)モデル [26] の関連型を二項間に限ったモデル=BERM [27] と等価なモデルとなった。また、EGDMの拡張データ定義および操作は、独立したフェーズでプリコンパイルされ、G-BASE上に実装されるので、概念設計から実装に至るまで、よりスムーズなデータベース設計が行えるようになった。

さらに、翌年の藤代による再考 [17] により、EGDMは、関係モデルの直接の拡張として知られている、Codd自身によるRM/T [28] やWiederholdらの構造(Structural)モデル [29] 等と同等以上のデータベース抽象化の記述能力を備えるようになった。これら3つのモデルは、基礎モデルのデータ構成子(constructor, 関係モデルならば関係スキーマ, GDMならば、レコード型, リンク型)の性質と役割をいくつかに分類し、それぞれに独自の呼称と一貫性制約を付加する方式で、意味データモデル [30] が直接扱っている汎化 [31] を始めとするデータベース抽象化を実現している。この意味から、RM/Tや構造モデルと同様、EGDMを疑似意味データモデルと呼ぶ[17]。以下で用いるEGDMは、文献 [17] の定義をファジィモデリング向けに再拡張したものである。

3.3 EGDMのデータ定義

EGDMのレコード型は、その役割に応じて

- a) 実体(E)-レコード型
- b) 属性(P)-レコード型

の二種類に大別される。通常、E-レコード型は、対応する実体型の識別子をフィールドに含む。また、P-レコード型に分離して扱うまでの必要はない属性もフィールドに追加してよい(例: 図3(b)のPERSONにおけるSEX)。一方、P-レコード型は記述の対象となる実体型の個々の属性ごとに用意され、属性値を保持する単一のフィールドからなる。

さらにP-レコード型は、可能性分布を効果的に表現するために

- b-1) 値属性(VP)-レコード型
- b-2) 記述属性(DP)-レコード型

の二種類に分けられる。なお、クリस्पな属性データはすべてVP-レコード型によって記述されるとする。

実体型と複数の属性との各関連型は、E-レコード型とP-レコード型との間の実リンク型、すなわち実体型の識別子と属性値との多対多の対によって表現される。このリンク型も、結びつくP-レコード型の種類に応じて、

- C-1) VP-リンク型
- C-2) DP-リンク型

の二種類に分けられる。また、VP-レコード型とDP-レコード型との間も、

- C-3) PP-リンク型

によって連結されるとする。

EGDMには、この他にもデータベース抽象化に対応する6種類のリンク型が用意されているが、ここでは論じない [17]。

3.4 ファジィ属性データのEGDM表現

各P-レコード型が

- ・DP-レコード型: 可能性分布ラベル
- ・VP-レコード型: 可能性分布の要素をそれぞれ格納すると仮定する。すると、各P-リンク型は
- ・VP-リンク型: 可能性をリンクフィールドとする可能性分布による属性
- ・DP-リンク型: 可能性分布ラベルによる属性
- ・PP-リンク型: 可能性をリンクフィールドとする可能性分布ラベルの定義

をそれぞれ表現することになる。FRDBで言うならば、PP-リンク型がEDBに格納され、その他のリンク型は、VDBに格納されるものと考えられる。式(3.2)からもわかるように、可能性は本来、実体とその属性の実現値の対から一意に決まる性質を持つので、EGDMを用いる場合、可能性をそれらの間のリンクフィールド値としてモデリングするのが極めて自然である。

この方針に沿って、図3(a)に示された関係スキーマをEGDMでモデリングしなおした結果を図3(b)に示す。また、図3(a)の実現値の例は図6(b)

に示すような形に格納される。両図では、“構成子名/”の直後にその分類名が指示されている。

さて、1項で述べた関係スキーマにおける非正規化と空値問題が、EGDMのリンク固有の表現によって効果的に解決されていることを以下に示そう。

まず、P-リンク型の実現値が多対多になることが許されているので、P-レコード型自身は正規化することができる。したがって、非正規形の問題は解消される。なお、可能性分布における可能性のデフォルト値は1とする。したがって、クリスプな年齢の場合には、1対1に年齢の実現値が対応し、その可能性は1となる。

空値“unknown”は、すべての年齢と関連する可能性があると解釈できるので、AGEに現れるすべての年齢の実現値とリンクしてもよいが、リンクの実現値確保のための空間効率と、unknownと判断するための全数検査による時間効率とを考慮して、VP-レコード型AGEの特別な実現値“null”を指すことにする。これは、馬野の可能性分布－関係モデル [8] で採用されている方針と基本的に同じである。また、“未定義(undefined)”の空値の場合は、年齢に関するリンク実現値が全くないことに相当する。これはリンク指向特有の表現方法である。

ここで、“適応不可能(inapplicable)”の空値は、EGDMの場合、その特化機構 [17] によって対応する属性をもたない別の実体型を表すレコード型が作られるので、考察対象から前もって除かれていることに注意されたい。

なお、リンク指向のファジィデータ表現として、馬野らによるファジィ集合論的データ構造 (FSTDS) [32] が既に知られている。しかし、FSTDSは計算機内部表現向きのデータ構造であり、EGDMのような汎用データモデルがとる実現値のリンク構造とは、対象とするデータの抽象度が本質的に異なっていると考えられる。

4. EGDMによる主観的ファジィ実現値ビュー

本節では、前節で述べたEGDMによるファジィ属性データ表現を基にして、2.2節で予め示しておいた主観的实现値ビューの概念を用いて、個別化されたファジィ属性データの統合的管理方式を提案する。

4.1 主観的实现値ビューの実現

EGDMを用いて主観的实现値ビューを実現するためには、基本的にはFDBSのデータ辞書内に登録されるシステムのユーザIDごとに、P-リンク型の実現値を管理すればよい。

概念モデリングの段階では、P-リンク型のリンクフィールドに、ユーザIDの指定を追加すれば十分である。このリンク型が主観的实现値ビューごとにその実現値が異なる特別なリンク型であることを表現するために、リンク型名の直前に“*”を付けることにする。この方針に沿って、図3(b)を書き改めた結果を図4に示す。

図4 主観的实现値ビュー構成のためのEGDMスキーマ

しかしFDBSの実装に当たっては、主記憶上にロードされるデータ量を考慮しなければならない。もし、ユーザIDがP-リンク属性に追加指定されたままシステムが実装されるならば、個別の問合せを評価する際に、大量の不必要な他人のリンク実現値も同時に主記憶上にロードされることになる。したがって、仮想記憶管理下ではページングの回数が圧倒的に増加し、時間効率を著しく劣化させることになる。

そこで、2.2節で示した方針に沿って、P-リンク

型の実現値リストを、主観的ファジィデータストレージに、概念データベースから独立させて保持・管理し、そこから個々のユーザが必要とするリンク実現値だけを取り出し、概念データベースと結合し、主観的実現値ビューを構成することが考えられる。この方針に従えば、主記憶の容量が増大し、主記憶データベースの有効性が論議されている現在の状況を考え合わせると、実用上十分な個別化されたファジィデータの統合的管理が実現できると考えられる。

4.2 FDBSの基本アーキテクチャと問合せ手続き

前項の方式に基づくFDBSの基本アーキテクチャを図5に示す。FDBSは、大きく6個の基本システムモジュールから構成される。データ辞書には、FDBSのメタデータが管理されている。具体的には、ユーザID等のユーザ管理情報を始め、概念データベースや主観的ファジィデータストレージに実際に表現されているレコード型やリンク型の名称、その役割名等も記述されている。データマネージャは、FDBSの管理プログラム群であり、シス

テム全体の制御を行うカーネルモジュールである。

ここで、図5のデータ・制御の流れに沿って、問合せ処理手続きの概略を述べる。

- ①ユーザはまず個々のIDとユーザ独自の問合せをシステムに入力する。
- ②ユーザIDと問合せがデータマネージャに渡され、その解釈が行われる。
- ③入力されたユーザIDが登録されているかどうかデータ辞書に照会される。正しくなければ、⑧⑨と進み、エラーメッセージをユーザに返す。
- ④正しい場合は、データマネージャにOKのメッセージが返る。
- ⑤それを受けて、予め解析してあった問合せ内容に応じて、処理に必要な範囲のサブデータベースを概念データベースから、該当するユーザのファジィリンクデータを主観的ファジィデータストレージからそれぞれロードするようなメッセージを、データマネージャが発行する。
- ⑥⑤のロードが実行され、主観的ファジィ実現値ビューが主記憶上で合成される。その上で、前述の問合せ処理が行われる。
- ⑦問合せの結果がデータマネージャに返される。
- ⑧さらに、その結果はインターフェイス内で表示用に修飾される。
- ⑨問合せ結果がユーザに返される。

- DD : データ辞書
 DM : データマネージャ
 I/F : インターフェイス
 SOV : 主観的実現値ビュー
 CDB : 概念データベース
 SFDS : 主観的ファジィデータストレージ
 → : データ・制御の流れ

図5 FDBSの基本アーキテクチャ

図5に示したアーキテクチャに基づいたFDBSは、G-BASE上に容易に実装可能である。実際には、既存のデータマネージャとデータ辞書の機能を補うためのサブDBMSと、インターフェイスに対応する応用プログラムを構築することになる。このとき、概念データベースだけでなく、サブデータ辞書、主観的ファジィデータストレージ、主観的ファジィ実現値ビューのいずれもEGDMによって記述可能であることに注意されたい。このことは、EGDMの記述能力の高さとG-BASEの優れた拡張性を示唆している。

4.3 主観的ファジィ実現値ビューと問合せの例

図 6(a)に、図 3(b)に示した年齢データベースのEGDMスキーマに対応する主観的ファジィデータストレージの実現値の例を示す。ここでは分かりやすいように表形式にして、ユーザA, B二人のファジィリンクデータが示されている。

さらに、これらから構成されるユーザA, Bの主観的ファジィ実現値ビューを図 6(b), (c)にそれぞれ示す。個々のユーザの年齢に対する主観的な見方の違いが現れているのがわかる。

以上から明らかなように、主観的ファジィ実現値ビューをEGDMを利用して実現する長所の一つは、個々のビュー構成のために、レコード実現値を修正することなく、対応するリンク実現値だけを主観的データストレージから抽出して置換すればよいことにある。この性質は、グラフ論的な表現のわかりやすさを単に得るだけでなく、効率的なデータ管理も可能にする一種のデータ独立性として捉えることができる。

最後に、この問合せ処理手続きによって、予め用意されていた個々のユーザのファジィ属性デー

years/VP

NAME	YEAR	USERID	m
Linda	18	A	0.5
Linda	19	A	0.9
Linda	20	A	0.8
James	17	B	0.3
James	18	B	0.4
Joseph	20	B	1.0

looks/DP

NAME	GROUP	USERID
James	young	A
Linda	young	B

breakdown/PP

GROUP	YEAR	USERID	m
young	17	A	0.7
young	18	A	1.0
young	19	A	0.9
adult	20	A	1.0
adult	40	A	1.0
middle	40	A	1.0
young	17	B	1.0
young	18	B	0.9
adult	20	B	0.8
middle	40	B	1.0

(a) 主観的ファジィデータストレージ

(b) ユーザAのビュー

(c) ユーザBのビュー

図 6 EGDM年齢データベースにおける主観的ファジィ実現値ビューの構成例

タの違いが、一定の対象世界を扱うデータベースに対して、同一の問合せを発行しても、異なる結果を生む効果を与えることを示す。例として、「0.5以上の可能性で若いと考えられる人は誰か」という問合せを評価しよう。問合せ言語は、EGDMに基づく操作言語EGDLを用いる。紙面の都合上、ここでは文法の詳細は割愛するが、英語に近い表現であるため、意味するところは容易に理解されよう。

図4のスキーマでは、可能性分布要素のレコード型AGEを経るかどうかで、以下のような二種類のアクセスパスが考えられる：

(a：ラベルを介さない場合)

```
for AGE_GROUP [GROUP="young"]
  ../inverse looks/.. PERSON
  print PERSON.NAME;
```

(b：ラベルを介す場合)

```
for AGE_GROUP [GROUP="young"]
  ../breakdown/.. AGE
  ../inverse years/.. PERSON
  [min(breakdown.m, years.m) >= 0.5]
  print PERSON.NAME;
```

上述の手続きから明らかなように、USER_IDに関する指示は問合せに先立ちシステムに入力されているので、主観的ファジィ実現値ビューの上の問合せ文にはUSER_IDの条件は含まれていないことに注意されたい。

それでは、図6にしたがって、上記の問合せ文を評価してみよう。まず、ユーザAが問合せた場合は、図6(b)の実現値により、

a : James b : Linda

という結果を得る。同様にユーザBの場合は、図6(c)より、

a : Linda b : ϕ (空集合)

という結果を得る。ユーザBのパスbの結果が空集合となるのは、Jamesの年齢の可能性が相対的に低く、どのアクセスパスを通っても、合成された可能性が0.5より低くなるためである。

最終結果は、両方のパスの結果の合併であるから、結局、同一の問合せに対する両者の評価は異なることになる。

5. おわりに

5.1 総括

本論文では、疑似意味データモデルEGDM上に定義された主観的実現値ビューという概念を用いて、個別化された可能性分布によって表現されたファジィ属性データを統合的に管理する方式を提案した。属性値として可能性分布だけを扱うという本論文の仮定は、主観的ファジィデータストレージの格納対象をリンク型データだけに限定するので、客観的な固定ノード集合に対し、主観的な重み付きアークの集合を選択して、ユーザごとに異なるファジィグラフを構成するという明解なビュー構成を実現することができた。これは、概念設計モデルとしてのEGDMの特長を最もよく表している。

また、より複雑なファジィ属性を扱う主観的ファジィ実現値ビューも考えられる。例えば、DP-リンク型のリンクフィールドに可能性を追加表現すれば、属性名にファジィ集合ラベルをもち、その帰属度の数値そのものを属性値とするファジィ属性 [3] [12] も同時に複数表現できる。

本論文では、検索のための問合せに限って述べてきたが、EGDMによる主観的実現値ビューは、概念スキーマの基本構成子以外の型を導出しないので、個別化されたファジィデータのビュー更新手続きに、関係モデルの結合演算に相当する面倒なリンク操作を必要としない。したがって、検索と同様に、主観的ファジィ実現値ビューの更新トランザクションも、EGDLによって容易に記述できると考えられる。この点も、主観的実現値ビューが従来型のビューと本質的に異なる点である。

さらに、主観的実現値ビューは一般的な概念であるため、2.2節で予備的に示した関係モデルのように、他のデータモデルおよびそれに基づくDBMSに対しても、適用可能であることに注意されたい。

5.2 今後の研究方針

現在、上述したビュー管理機能を持つパイロッ

トFDBSをG-BASE上に実装する計画を検討している。その際、本論文で提案したファジィデータ管理方式の有効性を保証するために、主観的ファジィ実現値ビュー実現時の主記憶管理の効率に関する定量的解析が不可欠である。

また、小高ら [33] は、原始ファジィ属性を集約化し、複合ファジィ属性を合成する際の評価関数として、通常のファジィ集合理論における最大・最小に、双対な調和平均、双対な幾何平均、算術平均を加え、計7種類の平均演算子を用意し、問合せ時に個々のユーザがその中から自由に選択し、検索の"くせ"を出せるような個別化を提案している。ファジィ化されたデータと共に、ファジィ化された検索ツールの統合的管理も今後のシステムの拡張対象である。

理論的な面では、適当なファジィデータを入力しておき、仮の問合せをシステムに発行して、その結果と自己の予想とのズレを修正するように、主観的なファジィデータをチューニングするFDBSのフィードバック自己改善、さらには、文献 [32] [33] 等で部分的に考察されている汎化等の、データベース抽象化機構のファジィ化も興味ある問題として残されている。

謝辞

まず、本論文を執筆するに当たり、熱心に討論を重ねて下さった、徳島大学の宮本定明教授、大阪大学の馬野元秀助教授、国際ファジィ工学研究所の深海悟博士を始めとするファジィ・データベース・ファジィ情報検索研究会各位、ならびにクボタコンピュータ(株)の茅暁陽博士に対し、深謝申し上げます。また、著者らが常々G-BASEを利用できるようご配慮いただいている(株)リコーソフトウェア事業部の國井秀子部長にお礼申し上げます。最後に、査読者の方々の貴重なコメントが、本論文の記述の修正に大いに役立ったことを記しておきます。

参 考 文 献

- [1] Zadeh, L.: "Fuzzy Sets," *Information and Control*, Vol.8, pp.338-353 (1965).
- [2] 馬野元秀: 「データベース」, ファジィシステム入門(寺野寿郎他共編), オーム社, 第14章(1987)
- [3] 馬野元秀: 「ファジィ・データベースの現状と動向」, アトバンスト・データベース・システム・シンポジウム'89 論文集, pp.207-214(1989)
- [4] Kunii, T. L.: "DATAPLAN: An Interface Generator for Data Semantics," *Information Sciences*, Vol.10, pp.279-298(1976).
- [5] Tahani, V.: "A Conceptual Framework for Fuzzy Query Processing - A Step Toward Very Intelligent Database Systems," *Information Processing and Management*, Vol.13, pp.289-303 (1977).
- [6] Zadeh, L.: "PRUF - A Meaning Representation Language for Natural Languages," *Int. Journal of Man-Machine Studies*, Vol.10, pp.395-460 (1978).
- [7] Buckles, B. P. and Petry F. E.: "A Fuzzy Representation of Data for Relational Databases," *Fuzzy Sets and Systems*, Vol.7, pp. 213-226 (1982).
- [8] Umamo, M.: "FREEDOM-0: A Fuzzy Database System," Gupta, M. M. et al (eds.) *Fuzzy Information and Decision Processes*, North-Holland, pp.339-347 (1982).
- [9] Umamo, M.: "Retrieval from Fuzzy Database by Fuzzy Relational Algebra," in *Proc. IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis*, pp.1-6(1983).
- [10] Prade, H. and Testemale, C.: "Generalizing Database Relational Algebra for the Treatment of Incomplete or Uncertain Information and Vague Queries," *Information Sciences*, Vol.34, pp.115-143 (1984).
- [11] Zemankova-Leech, M. and Kandel, A.: "Implementing Imprecision in Information Systems," *Information Sciences*, Vol.37, pp.107-141 (1985).

- [12] Zemankova-Leech, M. and Kandel, A. : *Fuzzy Relational Data Bases - A Key to Expert System*, Verlag TUV Rheinland GmbH (1984).
- [13] 藤代一成, 白井靖人 : 「ファジィデータを含む外延的ビューの実現」, 第2回ファジィ・データベース・ファジィ情報検索研究会報告書, 印刷中 (1990)
- [14] Codd, E. F. : "A Relational Model of Data for Large Shared Data Banks," *Comm. ACM* Vol. 13, pp.377-387 (1970).
- [15] Brodie, M. L., Mylopoulos, J. and Schmidt, J. W. (eds.) : *On Conceptual Modelling - Perspective from Artificial Intelligence, Databases and Programming Languages*, Springer-Verlag (1984).
- [16] Fujishiro, I., Shirai, Y., Kunii, H.S. and Kunii, T. L. : "Extending External Views on a Link-Oriented Data Model," in *Proc. IEEE COMP-SAC87*, Tokyo, pp.287-403 (1987).
- [17] Fujishiro, I. : *A Logical Design Methodology for Link-Oriented Databases Using Data Model Mapping*, Doctoral Thesis, Faculty of Science, The University of Tokyo (1988).
- [18] Shirai, Y., Fujishiro, I. and Kunii, T.L. : "Graph-Based Formulation of Fuzzy Data," Janko, W. H. et al. (eds.), *Progress in Fuzzy Sets and Systems*, Kluwer Academic Publishers, pp.147-160 (1990).
- [19] Shirai, Y., Fujishiro, I. and Kunii, T.L. : "An Entity-Relationship Approach to Fuzzy Data Modeling," *ファジィ・データベースおよびファジィ情報検索に関する研究報告書*, pp.47-69 (1990).
- [20] Zadeh, L. : "Fuzzy Sets as a Basis for a Theory of Possibility," *Fuzzy Sets and Systems*, Vol.1, pp.3-28 (1978).
- [21] Date, C. J. : *An Introduction to Database Systems*, 5th Edition, Addison-Wesley (1990).
- [22] Kunii, H. S. : *Graph Data Model and Its Data Language*, Springer-Verlag, 1990 [*Graph Data Language : A High Level Access-Path Oriented Language*, Ph.D. Thesis, The University of Texas at Austin, 1983].
- [23] 藤代一成, 國井秀子 : 「GDM-スーパーリレーショナルモデル」, コンピュートロール, コロナ社, Vol.17, pp.14-20 (1987)
- [24] (株) リコー (編) : G-BASEマニュアル (1987)
- [25] 國井利泰, 國井秀子 (共編) : 「スーパーリレーショナルDBMS」, コンピュートロール, コロナ社, Vol.17 (1987)
- [26] Chen, P. P. : "The Entity-Relationship Model-Toward a Unified View of Data," *ACM Trans. Database Systems*, Vol.1, No.1, pp.9-36 (1976).
- [27] Chen, P. P. : "A Preliminary Framework for Entity-Relationship Models," in P. P. Chen (ed.), *Entity-Relationship Approach to Information Modeling and Analysis*, North-Holland, pp. 19-28 (1983).
- [28] Codd, E. F. : "Extending the Database Relational Model to Capture More Meaning," *ACM Trans. Database Systems*, Vol.4, No.4, pp.397-434 (1979).
- [29] Wiederhold, G. : *Database Design*, 2nd Ed., McGraw-Hill (1983).
- [30] Hull, R. and King, R. : "Semantic Database Modeling : Survey, Applications and Research Issues," *ACM Computing Surveys*, Vol.19, No. 3, pp.201-260 (1987).
- [31] Smith, J. M. and Smith, D. C. P. : "Database Abstractions : Aggregation and Generalization," *ACM Trans. Database Systems*, Vol.2, No. 2, pp.105-133 (1977).
- [32] Umano, M, Mizumoto, M. and Tanaka, K. : "FSTDs System : A Fuzzy-Set Manipulation System," *Information Sciences*, Vol.14, pp.115-159 (1978).
- [33] 小高俊之, 阿倍博信, 横山光男, 松下温 : 「平均演算子を利用したあいまい要求処理法」, 情報処理学会データベース・システム研究会 76-2 (1990).

(1990年8月29日受付)

[問い合わせ先]
〒305 茨城県つくば市天王台 1-1-1
筑波大学電子・情報工学系
藤代一成
☎ : 0298-53-5344
☎ : 0298-53-5206
E-mail fujishiro@is.tsukuba.ac.jp

著者紹介

藤代 一成 (ふじしろ いっせい)

筑波大学 電子・情報工学系

1983年筑波大学第三学群情報学類卒業, 1985年筑波大学大学院博士課程工学研究科電子・情報工学専攻を修士取得退学後, 東京大学理学部情報科学科助手となる。1988年筑波大学電子・情報工学系に移り, 現在同講師, 理学博士(東京大学)。データ工学, 特に視覚/数値情報の表現, 管理とデータベース設計方法論に関する研究に従事。主な著書に, 「CAD/CAM」(編著, 丸善), *Link-Oriented Database Design* (共著, Springer-Verlag社, 執筆中) 等がある。情報処理学会, IEEE各会員。

池辺 八州彦 (いけべ やすひこ)

筑波大学 電子・情報工学系

1934年京都生まれ。京都大学工学部より学士, 修士号をそれぞれ1957年, 1959年に, 米国テキサス大学オースチン校より1966年に数学でPh.D.を授かる。京都産業大学(66-68年), 京都大学(68-69年), テキサス大学オースチン校(69-76年), 米国ノースウェスタン大学(76-78年)を経て, 現在筑波大学教授。数値解析および二ヶ国語ハイパー辞書に関する研究に従事。

白井 靖人 (しらい やすと)

静岡大学 教育学部

1984年トロント大学文理学部卒業(計算機科学専攻)。1987年東京大学大学院理学系研究科修士課程修了(情報科学専攻), 同年同大学院博士課程進学, 1989年同課程を退学し, 同大学助手。1990年静岡大学教育学部助教授。高度利用データベースシステム, 情報システム, コンピュータグラフィックスとその応用に関する研究に従事。ACM, IEEE, 情報処理学会, コンピュータ・グラフィックス学会, 日本ファジィ学会会員。

國井 利泰 (くにい としやす)

東京大学 理学部

学士, 修士, 博士号を化学分野で各々1962, 1964, 1967年に東京大学より受ける。1969年東京大学理学部助教授。1977年日本科学技術情報センター丹羽賞受賞。1978年東京大学理学部教授。現在に至るまでに執筆・編集した情報科学関係の単行本の数は32冊を越える。また, 審査済み学術論文も120余りを数える。1984年コンピュータ・グラフィックス学会(CGS)を創設, 1984年より1988年まで会長。学術誌「*The Visual Computer*」をSpringer-Verlag社より創刊(編集委員長)。The Journal of Visualization and Computer Animation (John-Wiley & Sons, Chichester)編集副委員長。IEEE Computer Graphics and Applications, IEEE Transactions on Knowledge and Data Engineering編集委員。IEEEフェロー。