

ファジィ情報検索システム INDAS/ff の試作

佐藤和洋¹ 田坂光伸¹ 山本洋一¹ 浪岡美予子¹ 茂木啓次²

1: (株) 日立製作所システム開発研究所 2: 日立マイクロコンピュータエンジニアリング(株)

我々は、人間の知的活動、そしてそれに伴う意思決定の曖昧な情報処理及び管理を効果的に支援するためのファジィ情報処理機能について検討を進めている。その中で本稿では、データベースに対する問合せの曖昧性を対象としたファジィ情報検索システム INDAS/ff について論じる。ファジィ問合せはSQLにファジィ述語を導入することで実現し、ファジィ問合せに対する新たな適合度評価方式を提案する。また、ユーザの情報取得過程のフォーカシングを効率的に支援するために、適合度分布に基づいた概略結果情報表現であるサマリボックスや、またその極め細かい適合度状態を表示する適合度マップ等からなる視覚的インターフェース VFN(View Focusing Navigator)を提示する。

The fuzzy information retrieval system INDAS/ff

Kazuhiro Satoh¹ Mitsunobu Tasaka¹ Yo-ichi Yamamoto¹ Miyoko Namioka¹ Keiji Moki²

1: Systems Development Laboratory, Hitachi,Ltd.

1099 OHZENJI ASAOKU, KAWASAKI-SHI, 215 JAPAN

2: Hitachi Microcomputer Engineering,Ltd.

5-22-1 JOSUIHON-CHO, KODAIRA-SHI, 187 JAPAN

We promote our study on the fuzzy information processing functions, which can efficiently support a 'fuzzy' information processing and management in human's intellectual activities, which contain decision making processes.

This article, involved in the study, gives an outline of the fuzzy information retrieval system INDAS/ff, whose objective is to cope with a 'fuzzy' query. In order to express the fuzzy query, SQL is extended, and a new method to evaluate an object's adaptability to the fuzzy query is proposed. The system also provides users with a visual interface, VFN(View Focusing Navigator), to give an effective decision aid to their view focusing in the information retrieval processes.

1 - はじめに

高機能、高性能且つ低価格な計算機資源及び分散環境の普及浸透により、マルチメディアデータを対象とした高度なデータ処理環境が実現されつつある。これに伴い、処理対象データのデータベース（以下DBと略）化と共に、それを管理するDBシステムの高機能化及び高性能化が必要となっている。^{10,11}

このような状況に鑑み、我々は現在の各種先端情報処理技術を融合させた新しいDBシステム（知的DBシステムINDAS（INtelligent DataBase System）という）を検討している。即ち、INDASは大量な情報を自然な形式で維持管理し、DBに存在する情報が既にそのユーザの記憶の中にあるかのように容易にその情報を利用でき、そのユーザの能力を支援代行及び増長する機構を装備したものを想定している。その中の一つとして、人間の知的活動、そしてそれに伴う意思決定の曖昧（以下ファジィと略）な情報処理及び管理を効果的に支援するためのファジィ情報処理機能の検討を進めている。ファジィな情報処理としては、問合せのファジィ性、情報そのもののファジィ性、情報処理過程のファジィ性等、種々のファジィ性を対象とする必要があり、これまでにもこれらを対象として、ファジィ情報検索、ファジィDB、不完全情報管理、視覚的ユーザインタフェース等、多くのファジィ情報処理に関する研究がなされている^{1~5,9}。昨今のファジィ理論の実用的な分野への適用及び普及により一層注目され、ファジィ理論を適用した上記テーマに関する研究が様々な研究機関で活発化しているが、解決すべき技術課題も多く、今後更に人間の情報処理により親和性のある情報処理システムの実現に向けて活発な研究開発が進められるものと思われる^{1,2,8~10}。

さて、本稿では上記ファジィ性の中の問合せのファジィ性を対象としたファジィ情報検索機構INDAS/fuzzy facilityの問合せ評価方式、システムアーキテクチャ、視覚的インタフェース等について論じる。まず第2章では、本稿に関係するファジィ集合論の概要及び本稿が対象とするファジィ情報検索に関する問題点について述べ、次に第3章では、新しく提案するファジィ問合せの適合度評価方式について論じ、続いて第4章では上記適合度評価方式を組んで試作したファジィ情報検索システムの構成、ファジィ問合せ言語及び視覚的インタフェースVFN

(View Focusing Navigator)等の機能概要について述べる。最後に本稿の締めとして今後の計画等について述べる。

2 - ファジィ集合論の概要と問題提起

2 - 1 ファジィ集合と基本演算

（1）ファジィ集合

【定義1】全体集合Ωにおけるファジィ集合Aとは、Ωの任意の要素ωがAに属する度合いを示す特性関数（メンバシップ関数と呼ぶ）μAによって

$$\mu A : \Omega \rightarrow [0, 1]$$

のように特性づけられた、ωの集まりとして定義される集合である。また、値μA(ω)は閉区間[0, 1]の実数値に対応し、要素ωのファジィ集合Aにおけるグレード（所属度）を表す。

ファジィ集合Aの表記法としては以下の方法がよく用いられる。まず、全体集合Uが有限集合{ωi | i=1, ..., n}の場合

$$A = \mu A(\omega_1)/\omega_1 + \mu A(\omega_2)/\omega_2 + \dots + \mu A(\omega_n)/\omega_n$$

$$= \sum_{i=1}^n \mu A(\omega_i)/\omega_i$$

ここで、/はセパレータで、+は結び(OR)を意味する。他の表記法もあるが本稿では上記記法を用いる。

（2）メンバシップ関数

メンバシップ関数としては種々の形態があるが、標準関数と呼ばれる下記の関数がよく使用される。これらの関数はその形状から各自S関数、Π関数、Z関数とも呼ばれる。

（i）S関数：S(ω; a, b, c) =

$$\begin{array}{ll} 0 & \omega \leq a \\ 2*((\omega-a)/(c-a)) & a \leq \omega \leq b \\ 1-2*((\omega-c)/(c-a)) & b \leq \omega \leq c \\ 1 & \omega \geq c \end{array}$$

（ii）Π関数：Π(ω; b, c) =

$$\begin{array}{ll} S(\omega; c-b, (2c-b)/2, c) & \omega \leq c \\ 1-S(\omega; c, (2c+b)/2, c+b) & \omega \geq c \end{array}$$

（iii）Z関数：Z(ω; a, b, c) =

$$1-S(\omega; a, b, c) \quad b=(a+c)/2$$

（3）ファジィ集合の基本演算

全体集合Uにおけるファジィ集合A, Bに関する基本的な演算としては次に示すようなものがある。

$$(a) 包含：A ⊆ B \Leftrightarrow \mu A(\omega) \leq \mu B(\omega), \forall \omega \in U$$

$$(b) 交集：A ∩ B \Leftrightarrow \mu A \cap B(\omega) = \mu A(\omega)T\mu B(\omega), \forall \omega \in U$$

$$(c) 和集：A ∪ B \Leftrightarrow \mu A ∪ B(\omega) = \mu A(\omega)S\mu B(\omega),$$

$\forall \mu$

(d) 様集合 : $\neg A \oplus \mu \neg A(\mu) = 1 - \mu A(\mu), \forall \mu$

ここで、Tは次の性質 T 1 ~ T 4

T 1 : T(0, 0)=0, T(x, 1)=T(1, x)=x (境界条件)

T 2 : T(x1, y1)≤T(x2, y2);

ここで、 $x_1 \leq x_2, y_1 \leq y_2$ (単調性)

T 3 : T(a, b)=T(b, a) (対称性)

T 4 : T(a, T(b, c))=T(T(a, b), c) (結合性)

を満たす関数 T : $[0, 1] \times [0, 1] \rightarrow [0, 1]$ で、 t - ノルムと呼ばれる積演算を表す。t - ノルムの例としては、以下のようなものがある:

$x \wedge y = \min(x, y)$ (論理積),

$x \cdot y = x \times y$ (代数積), 等。

また、Sは次の性質 S 1

S 1 : S(1, 1)=1, S(0, x)=S(x, 0)=x

及び上記 T 2 ~ T 4

を満たす関数 S : $[0, 1] \times [0, 1] \rightarrow [0, 1]$ で、 s - ノルム (あるいは t - コノルム) と呼ばれる和演算を表す。s - ノルムの例としては、以下のようなものがある:

$x \vee y = \max(x, y)$ (論理和),

$x + y = x + y - x \cdot y$ (代数和), 等。

なお、上記ノルムの間には、

$S(x, y) = 1 - T(1 - x, 1 - y)$

なる関係がある。

さらに、補集合は次の性質 N 1 及び N 2

N 1 : N(0)=1, N(1)=0

N 2 : N(N(x))=x (復帰性)

を満たす否定 N : $[0, 1] \rightarrow [0, 1]$ なる関数を用いて定義される。

上記の他にも、多くのファジィ集合演算子が提案され、所属度計算への適用が検討されている^{8, 14}。

2.2 本稿で扱うファジィ理論を用いた ファジィ情報検索に関する問題点

本節では、ファジィな問合せ例を元に、従来の適合度評価方式を示し、その問題点について論じる。

(1) ファジィ問合せの例

ファジィな意味合いの用語 (ファジィタームという) を用いた検索条件 (ファジィ検索条件といい、個々の述語をファジィ述語という) を含んだ問合せをファジィ問合せという。下記にその例を示す。

* 問合せ例 :

「年が若くて、給料の多い従業員を示せ」
この問合せをよく知られた DB 問合せ言語 SQL で表現すると、

SELECT *

FROM EMP

WHERE AGE is young AND SAL is high

となる。ここで、波線を付加した部分がファジィタームでこれらを含んだ

AGE is young AND SAL is high

がファジィ検索条件に対応する。

(2) 従来のファジィ問合せ適合度評価方式

従来上記のようなファジィ検索条件式 (ファジィ述語論理式ともいう) の扱いには前述したノルムを満足する種々の積演算及びその双対演算である和演算が適用されており、更に、平均演算子、補償演算子、自己双対演算子等様々なファジィ集合に対する演算子が数多く提案され、その適用が検討されている。これらの中で最も一般的に使用されているものは次に示す論理積と論理和を用いたものである。即ち、上記の問合せのような AND 結合 (Conjunction(AND:論理積)) の場合は、

$\text{MIN}[\mu \text{AGE-young}(x)/x, \mu \text{SAL-high}(y)/y]$

また、「若いか或いは給料の多い」のような OR 結合 (Disjunction(OR:論理和)) の場合は、

$\text{MAX}[\mu \text{AGE-young}(x)/x, \mu \text{SAL-high}(y)/y]$ のようにファジィ問合せの適合度 (満足度ともいう) 評価の計算が行われる。

(3) 従来の適合度評価方式の問題点

これまで主観性の演算系として複雑なファジィ集合演算子が数多く提案され、その妥当性が検討されているが未だ適当な演算子は見つかっていない。大規模なデータベースを対象とするという観点からは、如何に主観性を反映しているからといって、複雑なファジィ集合演算に多大な処理時間を要するのは問題であり、簡易な演算体系を採用することが望ましい。即ち、主観性をうまく反映した簡易な演算体系を検討する必要がある。このような観点から、現在主にファジィ情報検索システム或いはファジィ DB システムにおいて採用されているファジィ演算系は論理積及び論理和演算である。

しかしながら、上記のような論理積及び論理和演算を単純に適用すると、個々のファジィ述語のファジィグレード値 (以下適合度といふ) に相違があっても、ファジィ検索条件式全体の適合度が单一のファジィ述語の適合度に支配されてしまうという問題がある。各ファジィ述語の適合度間の差が大きい場合には特に問題となる。即ち、様々な角度からの問合せ条件に対して、最終結果が单一の適合度で抑えられることになり、ユーザの着眼点、注目点が捨象される心配があると同時に、取得情報の適合度の差異の不明確さによってユ

ーザのフォーカシング（情報取得のための焦点付け）のぼやけが生じ、ユーザの後続処理の負荷が増大することが懸念される。例えば、詳細情報への手探りアクセスの頻発、情報取得のための高負荷な試行錯誤の増大等の問題が発生すると予想される。

本稿では、上記のような問題点に対処するために、従来の論理積及び論理和演算体系をベースに各ファジィ述語の適合度の差異を考慮し、ユーザの心理的なファクタを加味した適合度評価方式を提案する。

(4) ファジィ情報検索におけるユーザインターフェースについて

ファジィ情報検索の結果は従来の検索結果とは異なり、適合度という新たな尺度が付加された情報集合である。ファジィ性の処理或いは管理方式についての研究は多いが、上記のようなファジィ情報処理を施して得られた結果集合に対する操作系、特にユーザインターフェースに関する議論は少ない。主観を伴うファジィ性を自動的に扱うには限界があり、効果的なファジィ情報検索システムを実現するためには、ファジィ情報処理結果を視覚的に表示し、その情報取得のための試行錯誤を効率的に支援するユーザインターフェースとの補完を考える必要がある。このような視覚的ユーザインターフェースにより情報取得の焦点化（フォーカシング）の促進と新しい焦点の誘発等、効率的なフォーカシング支援が期待できる。

上記のような観点から我々は、ファジィ情報検索を効果的に支援するために、問合せ適合度評価結果を視覚的に表示し、それに対する操作を可能とする視覚的インターフェースVNを提案する。

3. ファジィ問合せ適合度評価方式の提案

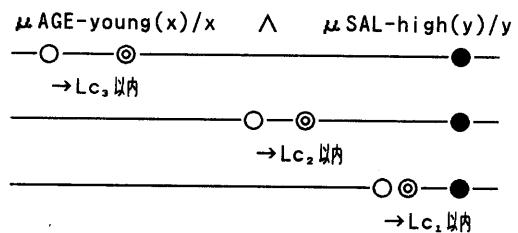
3.1 適合度評価方式の基本的な考え方

提案方式の基本的な考え方は、ファジィ述語間の適合度（ファジィグレード値）の差を考慮して、その差の大小によって、述語論理式全体の適合度を補正するものである。即ち、論理積の場合は全体の適合度を引き上げ補正（プラス補正）し、論理和の場合は引き下げ補正（マイナス補正）する方式である。そのための補正式をファジィ述語間の引力式と呼び、その補正值を引力値と呼ぶ。なお、補正值には上限を設けており、それを適合度の補正上限値或いは引力上限値（図中のLc等）と呼ぶ。以下に、簡単な論理積及び論理和を

例に図式的に上記考え方を示す。即ち、

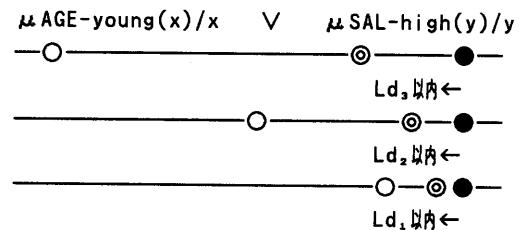
(i) Conjunction(AND結合)の場合

適合度の小さい方に対して、ファジィ述語間の適合度の差が小さい場合には小さい補正を、差が大きい場合には大きい補正を、各々引力上限値以内の範囲で補正しようとするものである。このことを図で表現すると下図のようになる。



(ii) Disjunction(OR結合)の場合

適合度の大きい方に対して、ファジィ述語間の適合度の差が小さい場合には小さい補正を、差が大きい場合には大きい補正を、各々引力上限値以内の範囲で補正しようとするものである。このことを図で表現すると下図のようになる：



3.2 適合度評価方式の詳細

はじめにConjunction及びDisjunctionについて記述し、次にこれらの標準形DNF及びCNFについて記述する。そして、これらに関して、まず簡便な適合度評価方式を示し、次に各述語間の引力式を考慮した評価方式を提示する。

(1) 簡易補正適合度評価方式

本方式は適合度評価処理負荷の低減を狙ったものであり、ファジィ述語論理式全体の中の適合度最大及び適合度最小のファジィ述語のみを対象に適合度評価を行うものである。なお、以下の記述において適合度評価範囲はmとし、 μ_i はファジィ述語*i*の適合度を、 $\sum \mu_i$ は (μ_1, \dots, μ_n) を意味し、 αc_j や αd_j は引力考慮範囲パラメタ、 a_{cj} や a_{dj} は引力重み付けパラメタ、そして Lc_j や Ld_j は

引力上限値を各々表すものとする。

(i) Conjunctionの場合

この場合、引力式と引力上限値との関係は以下のようになる：

- ① $\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i) \leq \alpha c_1$ の時
 $(\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)) * ac_1 \leq LC_1$
 . . .
- ② $\alpha c_j < \text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i) \leq \alpha ck$ の時
 $(\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)) * ac_j \leq LC_j$
 . . .
- ③ $\alpha c_l < \text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)$ の時
 $(\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)) * ac_l \leq LC_l$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようになる：

[AND 補正簡易適合度評価アルゴリズム]

```

if MAX(SUM μi) - MIN(SUM μi) ≤ αc1 then
    if (MAX(SUM μi) - MIN(SUM μi)) * ac1 ≤ LC1 then
        μc = MIN(SUM μi) + (MAX(SUM μi) - MIN(SUM μi)) *
            ac1;
    else
        μc = MIN(SUM μi) + LC1;
    . . .
else αcj < MAX(SUM μi) - MIN(SUM μi) ≤ αck then
    if (MAX(SUM μi) - MIN(SUM μi)) * acj ≤ LCj then
        μc = MIN(SUM μi) + (MAX(SUM μi) - MIN(SUM μi)) *
            acj;
    else
        μc = MIN(SUM μi) + LCj;
    . . .
else αcm < MAX(SUM μi) - MIN(SUM μi) then
    if (MAX(SUM μi) - MIN(SUM μi)) * acm ≤ LCm then
        μc = MIN(SUM μi) + (MAX(SUM μi) - MIN(SUM μi)) *
            acm;
    else
        μc = MIN(SUM μi) + LCj;
    μC = μc;

```

(ii) Disjunctionの場合

Conjunctionの場合と同様に、引力式と引力上限値は以下のようになる。

- ① $\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i) \leq \alpha d_1$ の時
 $(\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)) * ad_1 \leq LD_1$
 . . .
- ② $\alpha dj < \text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i) \leq \alpha dk$ の時
 $(\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)) * adj \leq LD_j$
 . . .
- ③ $\alpha dm < \text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)$ の時
 $(\text{MAX}(\sum \mu_i) - \text{MIN}(\sum \mu_i)) * adm \leq LD_m$

よって、上記のような引力補正を考慮した適合度評価式は以下のようになる：

[OR 補正簡易適合度評価アルゴリズム]

```

if MAX(SUM μi) - MIN(SUM μi) ≤ αd1 then
    if (MAX(SUM μi) - MIN(SUM μi)) * ad1 ≤ LD1 then
        μd = MAX(SUM μi) - (MAX(SUM μi) - MIN(SUM μi)) *
            ad1;
    else
        μd = MAX(SUM μi) - LD1;
    . . .
else αdj < MAX(SUM μi) - MIN(SUM μi) ≤ αdk then
    if (MAX(SUM μi) - MIN(SUM μi)) * adj ≤ LDj then
        μd = MAX(SUM μi) - (MAX(SUM μi) - MIN(SUM μi)) *
            adj;
    else
        μd = MAX(SUM μi) - LDj;
    . . .
else αdm < MAX(SUM μi) - MIN(SUM μi) then
    if (MAX(SUM μi) - MIN(SUM μi)) * adm ≤ LDm then
        μd = MAX(SUM μi) - (MAX(SUM μi) - MIN(SUM μi)) *
            adm;
    else
        μd = MAX(SUM μi) - LDm;
    μD = μd;

```

(2) 述語間補正適合度評価方式

次に、各ファジィ述語間の引力式を利用した適合度評価方式について記述する。また、簡単のために μP_i 等はファジィ述語 P_i の適合度を意味し、 αj や βj は引力考慮範囲パラメタ、 a_j や b_j は引力重み付けパラメタ、そして L_{1j} や L_{2j} は引力上限値を各々表すものとする。

(i) Conjunctionの場合

この場合、引力式と引力上限値との関係は以下のようになる。即ち、

- ① $|\mu P_1 - \mu P_i| \leq \alpha_1$ の時
 $|\mu P_1 - \mu P_i| * a_1 \leq L_{1,i}$
 . . .
- ② $\alpha j < |\mu P_1 - \mu P_i| \leq \alpha k$ の時
 $|\mu P_1 - \mu P_i| * a_j \leq L_{1,j}$
 . . .
- ③ $\alpha m < |\mu P_1 - \mu P_i|$ の時
 $|\mu P_1 - \mu P_i| * a_m \leq L_{1,m}$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようになる。

[AND 補正述語間適合度評価アルゴリズム]

```

for i=2 to n
    if |μP1 - μPi| ≤ α1, then
        if |μP1 - μPi| * a1 < L1,i, then
            μc = MIN[μP1, μPi] + |μP1 - μPi| * a1;
        else

```

```

 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + L_{1,i};$ 
    ...
else  $\alpha_j < |\mu P_1 - \mu P_i| \leq \alpha_k$  then
if  $|\mu P_1 - \mu P_i| * a_j < L_{1,j}$  then
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + |\mu P_1 - \mu P_i| * a_j;$ 
else
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + L_{1,j};$ 
    ...
else  $\alpha_m < |\mu P_1 - \mu P_i|$  then
if  $|\mu P_1 - \mu P_i| * a_m < L_{1,m}$  then
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + |\mu P_1 - \mu P_i| * a_m;$ 
else
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + L_{1,m};$ 
 $\mu P_1 = \mu c;$ 
each i

```

(ii) Disjunctionの場合

(i) の場合と同様に、引力式と引力上限値は以下のようにになる。即ち、

① $|\mu P_1 - \mu P_i| \leq \beta_1$ の時

$$|\mu P_1 - \mu P_i| * b_1 \leq L_{2,i}$$
 ...

② $\beta_j < |\mu P_1 - \mu P_i| \leq \beta_k$ の時

$$|\mu P_1 - \mu P_i| * b_j \leq L_{2,j}$$
 ...

③ $\beta_m < |\mu P_1 - \mu P_i|$ の時

$$|\mu P_1 - \mu P_i| * b_m \leq L_{2,m}$$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようにになる。
[OR結合補正適合度評価アルゴリズム]

```

for i=2 to n
if  $|\mu P_1 - \mu P_i| \leq \beta_1$  then
if  $|\mu P_1 - \mu P_i| * b_1 < L_{2,i}$  then
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - |\mu P_1 - \mu P_i| * b_1;$ 
else
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - L_{2,i};$ 
    ...
else  $\beta_j < |\mu P_1 - \mu P_i| \leq \beta_k$  then
if  $|\mu P_1 - \mu P_i| * b_j < L_{2,j}$  then
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - |\mu P_1 - \mu P_i| * b_j;$ 
else
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - L_{2,j};$ 
    ...
else  $\beta_m < |\mu P_1 - \mu P_i|$  then
if  $|\mu P_1 - \mu P_i| * b_m < L_{2,m}$  then
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - |\mu P_1 - \mu P_i| * b_m;$ 
else
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - L_{2,m};$ 
 $\mu P_1 = \mu d;$ 
each i

```

```

 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + L_{1,i};$ 
    ...
else  $\alpha_j < |\mu P_1 - \mu P_i| \leq \alpha_k$  then
if  $|\mu P_1 - \mu P_i| * a_j < L_{1,j}$  then
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + |\mu P_1 - \mu P_i| * a_j;$ 
else
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + L_{1,j};$ 
    ...
else  $\alpha_m < |\mu P_1 - \mu P_i|$  then
if  $|\mu P_1 - \mu P_i| * a_m < L_{1,m}$  then
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + |\mu P_1 - \mu P_i| * a_m;$ 
else
 $\mu c = \text{MIN}[\mu P_1, \mu P_i] + L_{1,m};$ 
 $\mu P_1 = \mu c;$ 
each i

```

(ii) Disjunctionの場合

(i) の場合と同様に、引力式と引力上限値は以下のようにになる。即ち、

① $|\mu P_1 - \mu P_i| \leq \beta_1$ の時

$$|\mu P_1 - \mu P_i| * b_1 \leq L_{2,i}$$
 ...

② $\beta_j < |\mu P_1 - \mu P_i| \leq \beta_k$ の時

$$|\mu P_1 - \mu P_i| * b_j \leq L_{2,j}$$
 ...

③ $\beta_m < |\mu P_1 - \mu P_i|$ の時

$$|\mu P_1 - \mu P_i| * b_m \leq L_{2,m}$$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようにになる。
[OR結合補正適合度評価アルゴリズム]

```

for i=2 to n
if  $|\mu P_1 - \mu P_i| \leq \beta_1$  then
if  $|\mu P_1 - \mu P_i| * b_1 < L_{2,i}$  then
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - |\mu P_1 - \mu P_i| * b_1;$ 
else
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - L_{2,i};$ 
    ...
else  $\beta_j < |\mu P_1 - \mu P_i| \leq \beta_k$  then
if  $|\mu P_1 - \mu P_i| * b_j < L_{2,j}$  then
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - |\mu P_1 - \mu P_i| * b_j;$ 
else
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - L_{2,j};$ 
    ...
else  $\beta_m < |\mu P_1 - \mu P_i|$  then
if  $|\mu P_1 - \mu P_i| * b_m < L_{2,m}$  then
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - |\mu P_1 - \mu P_i| * b_m;$ 
else
 $\mu d = \text{MAX}[\mu P_1, \mu P_i] - L_{2,m};$ 
 $\mu P_1 = \mu d;$ 
each i

```

(3) 一般論理式DNF及びCNFへの拡張
(1) 簡易補正適合度評価方式のDNF及びCNFへの適用

(i) DNFの場合

まずDNFの場合は、Conjunction数をx、上記(1)の(i)を用いて計算された論理積q(Conjunction q)の適合度を μC_q 、即ち、 $\mu C_q = \text{MIN}((\sum \mu_i)q) + (\text{MAX}((\sum \mu_i)q) - \text{MIN}((\sum \mu_i)q)) * \psi(ac)$

とすると、DNFの簡易補正適合度評価式は以下のようにになる。なお、各パラメタは上述のものと同様である。

まず、Conjunction間での引力式と引力上限値との関係は以下のようにになる。

① $\text{MAX}(\sum \mu C_q) - \text{MIN}(\sum \mu C_q) \leq \alpha c_1'$ の時
 $(\text{MAX}(\sum \mu C_q) - \text{MIN}(\sum \mu C_q)) * ac_1' \leq LC_1'$
...

② $\alpha c_j' < \text{MAX}(\sum \mu C_q) - \text{MIN}(\sum \mu C_q) \leq \alpha c_k'$ の時
 $(\text{MAX}(\sum \mu C_q) - \text{MIN}(\sum \mu C_q)) * ac_j' \leq LC_j'$
...

③ $\alpha c_m' < \text{MAX}(\sum \mu C_q) - \text{MIN}(\sum \mu C_q)$ の時
 $(\text{MAX}(\sum \mu C_q) - \text{MIN}(\sum \mu C_q)) * ac_m' \leq LC_m'$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようになる：

```
□ DNF簡易補正適合度評価アルゴリズム
if MAX(SUM μ Cq) - MIN(SUM μ Cq) ≤ α c1' then
    if (MAX(SUM μ Cq) - MIN(SUM μ Cq)) * ac1' ≤ LC1'
        then
            μ D = MAX(SUM μ Cq) - (MAX(SUM μ Cq) - MIN(SUM μ Cq))
            * ac1';
    else
        μ D = MAX(SUM μ Cq) - LC1';
    ...
else α c j' < MAX(SUM μ Cq) - MIN(SUM μ Cq) ≤ α c k'
    then
        if (MAX(SUM μ Cq) - MIN(SUM μ Cq)) * acj' ≤ LCj'
            then
                μ D = MAX(SUM μ Cq) - (MAX(SUM μ Cq) - MIN(SUM μ Cq))
                * acj';
        else
            μ D = MAX(SUM μ Cq) - LCj';
    ...
else α cm' < MAX(SUM μ Cq) - MIN(SUM μ Cq) then
    if (MAX(SUM μ Cq) - MIN(SUM μ Cq)) * acm' ≤ LCM'
        then
            μ D = MAX(SUM μ Cq) - (MAX(SUM μ Cq) - MIN(SUM μ Cq))
            * acm';
    else
        μ D = MAX(SUM μ Cq) - LCM';

```

(ii) CNFの場合

次にCNFの場合は、Disjunction数をy、上記(3)-1の(ii)を用いて計算された論理和r(Disjunction r)の適合度を μD_r 、即ち、

$$\mu D_r = \text{MAX}((\sum \mu_i)r) - (\text{MAX}((\sum \mu_i)r) - \text{MIN}((\sum \mu_i)r)) * \psi(ad)$$

とすると、CNFの簡易補正適合度評価式は以下のようにになる。

まず、Disjunction間での引力式と引力上限値との関係は以下のようにになる。

① $\text{MAX}(\sum \mu D_r) - \text{MIN}(\sum \mu D_r) \leq \alpha d_1'$ の時
 $(\text{MAX}(\sum \mu D_r) - \text{MIN}(\sum \mu D_r)) * ad_1' \leq LD_1'$
...

② $\alpha d_j' < \text{MAX}(\sum \mu D_r) - \text{MIN}(\sum \mu D_r) \leq \alpha d_k'$ の時
 $(\text{MAX}(\sum \mu D_r) - \text{MIN}(\sum \mu D_r)) * adj' \leq LDj'$
...

③ $\alpha dm' < \text{MAX}(\sum \mu D_r) - \text{MIN}(\sum \mu D_r)$ の時
 $(\text{MAX}(\sum \mu D_r) - \text{MIN}(\sum \mu D_r)) * adm' \leq LDm'$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようにになる：

```
□ CNF簡易補正適合度評価アルゴリズム
if MAX(SUM μ Dq) - MIN(SUM μ Dq) ≤ α d1' then
    if (MAX(SUM μ Dq) - MIN(SUM μ Dq)) * ad1' ≤ LD1'
        then
            μ C = MIN(SUM μ Dq) + (MAX(SUM μ Dq) - MIN(SUM μ Dq))
            * ad1';
    else
        μ C = MIN(SUM μ Dq) + LD1';
    ...
else α dj' < MAX(SUM μ Dq) - MIN(SUM μ Dq) ≤ α dk'
    then
        if (MAX(SUM μ Dq) - MIN(SUM μ Dq)) * adj' ≤ LDj'
            then
                μ C = MIN(SUM μ Dq) + (MAX(SUM μ Dq) - MIN(SUM μ Dq))
                * adj';
        else
            μ C = MIN(SUM μ Dq) + LDj';
    ...
else α dm' < MAX(SUM μ Dq) - MIN(SUM μ Dq)
    then
        if (MAX(SUM μ Dq) - MIN(SUM μ Dq)) * adm' ≤ LDm'
            then
                μ C = MIN(SUM μ Dq) + (MAX(SUM μ Dq) - MIN(SUM μ Dq))
                * adm';
        else
            μ C = MIN(SUM μ Dq) + LDm';

```

(ii) 詞語間補正適合度評価方式のDNF及びCNFへの適用

(i) D N F の場合

まず D N F の場合は、Conjunction数を \times 、上記 (2) の(i)を用いて計算された A N D 結合 q (Conjunction q)の適合度を μ_{Cq} 、即ち、

$$\mu_{Cq} = f_g / cq$$

とすると、D N F の適合度は以下のようになる。まず、Conjunction間での引力式と引力上限値との関係は以下のようになる。

$$① |fg/c1-fg/ci| \leq \alpha_1' \text{ の時}$$

$$|fg/c1-fg/ci| * a_1' \leq L_{11}'$$

...

$$② \alpha_{j'} < |fg/c1-fg/ci| \leq \alpha_k' \text{ の時}$$

$$|fg/c1-fg/ci| * a_{j'} \leq L_{1j'}$$

...

$$③ \alpha_m' < |fg/c1-fg/ci| \text{ の時}$$

$$|fg/c1-fg/ci| * a_m \leq L_{1m'}$$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようになる。
[DNF述語間補正適合度評価アルゴリズム]

```
for i=2 to x
    if |fg/c1-fg/ci| \leq \alpha_1' then
        if |fg/c1-fg/ci| * a_1' < L_{11}' then
            fg/c=MAX[fg/c1, fg/ci]-|fg/c1-
                fg/ci| * a_1';
        else
            fg/c=MAX[fg/c1, fg/ci]-L_{11}'';
    else
        if |fg/c1-fg/ci| * a_{j'} < L_{1j'} then
            fg/c=MAX[fg/c1, fg/ci]-|fg/c1-
                fg/ci| * a_{j'}';
        else
            fg/c=MAX[fg/c1, fg/ci]-L_{1j'};

    else
        if |fg/c1-fg/ci| * a_m < L_{1m'} then
            fg/c=MAX[fg/c1, fg/ci]-|fg/c1-
                fg/ci| * a_m';
        else
            fg/c=MAX[fg/c1, fg/ci]-L_{1m'};

    each i
```

(ii) C N F の場合

次に C N F の場合は、Disjunction数を y 、上記 (2) の(ii)を用いて計算された論理和 r (Disjunction r)の適合度を μ_{Dq} 、即ち、

$$\mu_{Dq} = f_g / dq$$

とすると、C N F の適合度は以下のようになる。まず、Disjunction間での引力式と引力

上限値との関係は以下のようになる。

$$① |fg/c1-fg/ci| \leq \beta_1' \text{ の時}$$

$$|fg/c1-fg/ci| * b_1' \leq L_{21}'$$

...

$$② \beta_{j'} < |fg/c1-fg/ci| \leq \beta_k' \text{ の時}$$

$$|fg/c1-fg/ci| * b_{j'} \leq L_{2j'}$$

...

$$③ \beta_m' < |fg/c1-fg/ci| \text{ の時}$$

$$|fg/c1-fg/ci| * b_m \leq L_{2m'}$$

よって、上記のような引力補正を考慮した適合度評価アルゴリズムは以下のようになる。

[CNF述語間補正適合度評価アルゴリズム]

```
for i=2 to y
    if |fg/d1-fg/di| \leq \beta_1' then
        if |fg/d1-fg/di| * b_1' < L_{21}' then
            fg/d=MIN[fg/d1, fg/di]+|fg/d1-
                fg/di| * b_1';
        else
            fg/d=MIN[fg/d1, fg/di]+L_{21}'';
    else
        if |fg/d1-fg/di| * b_{j'} < L_{2j'} then
            fg/d=MIN[fg/d1, fg/di]+|fg/d1-
                fg/di| * b_{j'}';
        else
            fg/d=MIN[fg/d1, fg/di]+L_{2j''};

    else
        if |fg/d1-fg/di| * b_m' < L_{2m'} then
            fg/d=MIN[fg/d1, fg/di]+|fg/d1-
                fg/di| * b_m';
        else
            fg/d=MIN[fg/d1, fg/di]+L_{2m''};
    fg/dl=fg/d;
each i
```

4 - ファジィ情報検索システムの構成と機能概要

4 - 1 システム構成

図1に試作したファジィ情報検索システムの構成を示す。当該システムは大きく、ビジュアルデータ処理部とファジィ問合せ処理部からなる。また、前者はビジュアルインタフェース部、ファジィ情報エディタ部、及びファジィ問合せ言語インタフェース部からなり、後者はファジィ問合せ処理制御部、ファジィ問合せ解析・変換処理部、ファジィ問合せ実行&評価処理部、及びファジィDD/D管理部等からなる。

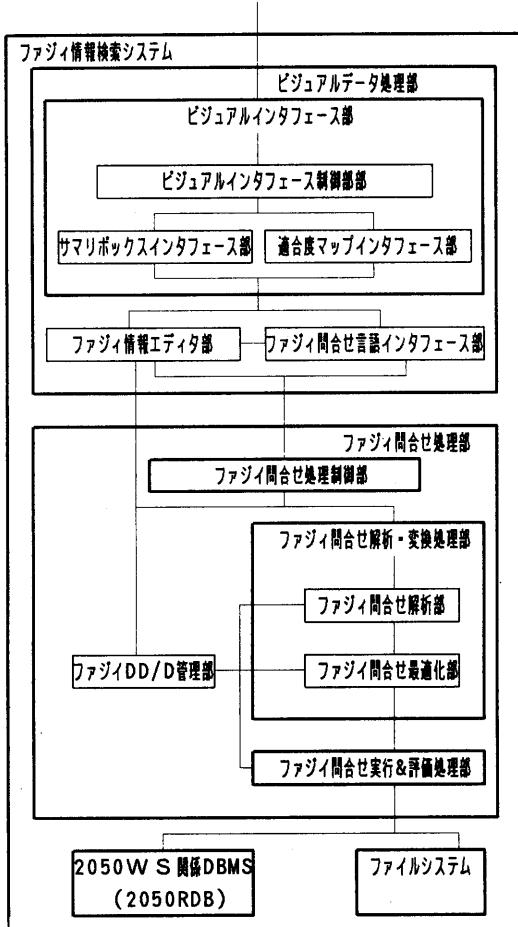


図1 実験システムの構成

当該システムの動作概略は以下の通りである。即ち、ビジュアルデータ処理部は、ファジィ問合せ言語による情報検索指示を受理し、それをファジィ問合せ処理部に渡す。ファジィ問合せ処理部は当該情報検索要求を実行し、取得データに検索条件に対する適合度を付加した形態で、その処理結果をビジュアルデータ処理部に渡す。ビジュアルデータ処理部のビジュアルインタフェース部は当該処理結果を後述するような視覚的な形態で表示し、引き続いて指定される情報取得表示操作指示に従い処理を行う。

4.2 機能概要

(1) ビジュアルデータ処理部の機能概要
ビジュアルインタフェース部は情報検索指示及び取得結果の視覚的操作を司る部分である。ファジィ情報エディタ部はファジィ D D / D で管理するファジィ関連情報の定義及び

編集操作を司り、ファジィ問合せ言語インターフェース部は上記ビジュアル及びエディティングインターフェースを、後述するファジィ問合せ言語インターフェースに変換し、処理結果を取得管理する。詳細は後述する。

(2) ファジィ問合せ処理部の機能概要

ファジィ問合せ処理部は問合せ処理システム全体を管理し、ファジィ問合せ解析・変換処理部のファジィ問合せ解析部はファジィ問合せの構文・意味解析を行い、ファジィ D D / D をアクセスし、後述するファジィ述語を含むファジィ問合せを単一テーブル（あるいはファイル）ファジィ部分問合せ（群）に展開し、当該ファジィ部分問合せ（群）から非ファジィ部分問合せ（群）を作成し、ファジィ問合せ最適化部は上記非ファジィ部分問合せ（群）の最適化処理を司る。また、ファジィ問合せ実行 & 評価処理部は上記非ファジィ部分問合せ（群）に基づいてデータベース或いはファイルアクセス用コマンドを作成発行し、その取得結果に対して適合度計算等を施し、適合度を付加した形態で問合せ言語インターフェース部に結果を渡す。また、ファジィ D D / D 管理部はファジィ項目、メンバシップ関数、ファジィ集合等の情報を管理する部分である。

4.3 ファジィ問合せ言語

ファジィ問合せ言語としては国際標準データベース言語SQLをベースに、下記形式のファジィ述語を新たに導入している：

```

<FUZZY述語> ::= </パリュー式> is [NOT]
                  <FUZZY関係子></パリュー式>
                  | </パリュー式> is [NOT] [<FUZZY修飾子>]
                  <FUZZY量子>
<FUZZY関係子> ::= ABOUT | NEAR | SIMILAR_TO | ...
<FUZZY修飾子> ::= VERY | MORE | MOST | ...
<FUZZY量子> ::= HIGH | LOW | SMALL | LARGE | ...
[問合せ例]
  { 價格が非常に安く、建物面積が約80m2の物件を示せ }
SELECT * FROM ESTATE, ARCH
  WHERE ESTATE.EPRICE is very low
    AND ARCH.AC_SIZE is about 80
  
```

4.4 視覚インターフェース VFN の概要

(1) 視覚インターフェースの狙い

本実験システムの視覚インターフェース VFN (View Focusing Navigator)は、ユーザ視点及びその視点の動きに対する焦点付け（フォーカシング）を実現し、ユーザの意思決定を効果的に支援するために、下記の観点に基

づいて開発を推進しているものである：

- (1)情報取得操作用多レベル簡易インタフェースの提供
- (2)取得情報のマクロ／ミクロ的視覚化表示フィードバック操作系の提供
- (3)ハイパーメディア風インタフェースの提供

(2) VFNの概要

視覚インタフェースVFNは複数のサブインターフェースの複合体である。その中に後述する、問合せ言語インターフェース、サマリボックスインターフェース、適合度マップインターフェース等がある。ユーザはまず、問合せ言語インターフェースを介して、ファジィ述語を含んだ問合せを入力する。そして、その実行結果がサマリボックスインターフェースにおける適合度に従ったサマリボックスとして表示される。当該サマリボックスに対して、更に詳細な適合度状態を必要とする場合には適合度マップインターフェースを介して、その詳細を知ることができる。以下では、これらについて詳述する。

(3) サマリボックスインターフェース

サマリボックスは問合せの処理結果をファジィ述語の適合度評価結果に従いマクロ的且つ視覚的に分類表示（概略結果表示）したものである。今、下記のような問合せが入力されたものとする：

/*問合せ例*/「価格が安くて、建物面積が約80m²の物件は？」

/*ファジィ問合せ表現*/

```
SELECT * FROM ESTATE, ARCH
```

```
WHERE ESTATE.EPRICE is low  
AND ARCH.ACsize is about 80
```

当該問合せの処理結果は図2に示すように、適合度分布に基づいた形式で概略表示される。これがサマリボックスで、該当するデータの量はボックスの厚みで表現され、ボックス内のデータはその適合度の降順にソートされている。当該ボックスに対して種々の操作を行うことができ、例えば、ボックスの表面をクリックして表示メニュー項目をクリックすると、クリックしたボックスの最も適合度の高いデータの詳細情報が表示され、ボックスの厚みの部分をクリックすると、クリックした部分に該当する適合度のデータが表示される。また、表示された個所の適当な部分をクリックすることにより、更に詳細な情報を取得することもできる。その結果を図3に示す。

概略結果表示

（サマリボックス）

```
問合せ: SELECT * FROM ESTATE, ARCH  
WHERE ESTATE.EPRICE is low  
AND ARCH.ACsize is about 80
```

問合せ適合度分布表示

100%	100-75%	75-50%
(0)	(12)	(22)

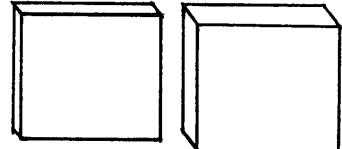


図2 サマリボックス表示例

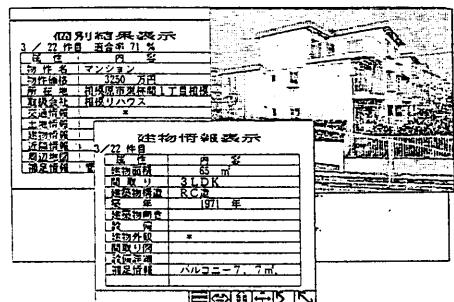


図3 詳細情報表示例

(4) 適合度マップインターフェース

適合度マップは、サマリボックスを範囲指定でクリックし、メニュー項目マップをクリックすることにより表示される。当該適合度インターフェースは、前述のサマリボックスインターフェースを補足するものである。即ち、サマリボックスは視覚的に大まかな判断を支援し、適合度インターフェースはサマリボックスで捨象された個別のファジィ述語の適合度を視覚的に表示し、ユーザの情報取得のためのフォーカシングを支援するものである。適合度マップ表示において、マーク表示が稠密状態にあって判読が困難な場合には、その部分を拡大表示させたり、また、適合度が同一のためマーク表示が重なっている場合には、その部分をクリックすると重なっているマークの種類をサブウインドウの形式で表示する。この例を図4に示す。また、この図の適当なマークをクリックすることにより、前述のサ

マリボックスインターフェースと同様、詳細情報を取得することができる。適合度マップ表示を終了したり、キャンセルすると前述のスマリボックス表示状態にもどる。

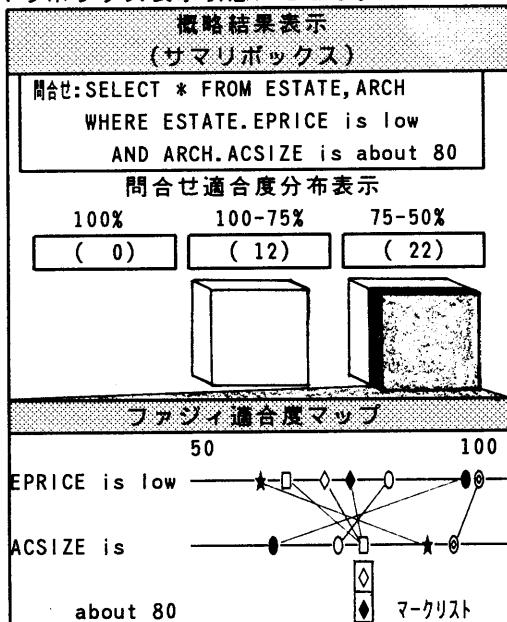


図4 適合度マップ表示例

上記の他、VFNの特徴的な機能の一つとしてオブジェクトホルダ機能がある。これは上記インターフェースを介して取得した情報を、ユーザの着眼点をインデックスとして保持する機能である。

以上のように、マクロからミクロ（或いはその逆も）という二段階アプローチと中間結果の情報取得管理等を視覚的に支援することにより、効果的なユーザのフォーカシングが期待できる。

5. おわりに

ファジィ情報検索実験システムのシステム構成、問合せ適合度評価方式及び視覚的インターフェースについて報告した。ファジィ述語をSQLに導入し、その適合度評価に関し新たな方式を提案すると共に、視覚的インターフェースに関しては、ユーザの情報取得過程のフォーカシングを効果的に支援するために、適合度分布に基づいた概略情報表現であるスマリボックス、及びその極め細かい適合度状態を表示する適合度マップ等のインターフェースを提示した。本実験システムはC言語で記述されており、日立のワークステーション2050/32上で動作する。

謝辞

本研究を推進するに当たり、御討論及び御支援頂いている関係各位に感謝致します。

参考文献

- 1) Tahani, Vo: A Conceptual Framework for Fuzzy Query Processing - A Step toward very Intelligent Database Systems, Information Processing & Management, 13, pp. 289-303(1977)
- 2) Lipski, w.: On Databases with Incomplete Information, J. ACM, 28, pp. 41-70(1981)
- 3) Prade, H., et al: Generalizing Database Relational Algebra for the Treatment of Incomplete or Uncertain Information and Vague Queries, Information Sciences, 34, pp. 115-143(1983)
- 4) Zemankova-Leech, M., et al: Fuzzy Relational Data Bases - A Key to Expert Systems, TUV Rheialand(1984)
- 5) Buckles, B. P., et al: Uncertainty models in information and database systems, J. of Information Sciences, 11, pp. 77-87(1985)
- 10) 木本: 最近のファジイ理論, 情報処理学会誌, 29, 1, pp. 11-22(1988.1)
- 6) 小橋: 決定を支援する認知科学選書 第18巻, 東京大学出版会(1988.6)
- 7) Motro, A.: VAGUE: A User Interface to Relational Databases that Permits Vague Queries, ACM TOIS, 6, 3, pp. 189-214(1988.7)
- 8) 高橋: リレーショナルDBのファジイ問合せ言語, 情報処理学会DBシステム研究会報告, 88-DBS-66(1988.7.22)
- 9) 佐藤: オブジェクト指向アプローチによるマルチメディアDBの研究動向, 情報処理学会マルチメディア通信と分散処理システムシンポジウム, pp. 21-30(1988.10)
- 11) 佐藤他: データベース技術 - CIM-DBシステムに向けて-, 日立評論 CIMシステム特集, 71, 6, pp. 9-15(1989.6)
- 12) Special Issues on Fuzzy Databases, Information Systems, 14, 6(1989.6)
- 13) 森田他: ファジイ文書検索システム(1)~実験システムと評価~, 情報処理学会第39回全国大会, 2N-2, pp. 1067-1068(1989.10)
- 14) 田中: ファジイ理論の過去・現在・未来, 日本ファジイ学会誌, 1, 1, pp. 48-62(1989.11)
- 15) 野村他: あいまい要求のための検索システム, 情報処理学会第40回全国大会, 1H-3, pp. 775-776(1990.3)
- 16) 平岩他: 抽象演算子のあいまい要求処理への適用, 情報処理学会第40回全国大会, 1H-4, pp. 777-778(1990.3)
- 17) 佐藤他: ファジイ情報検索システムの開発(1), 情報処理学会第40回全国大会, 3S-1, pp. 1081-1082(1990.3)
- 18) 田中他: ファジイ情報検索システムの開発(2), 情報処理学会第40回全国大会, 3S-2, pp. 1083-1084(1990.3)