

曖昧なデータを取り扱うデータベースの一構築法[†]杉原周治^{††} 宮崎 淳^{††} 相磯秀夫^{††}

近年、高機能なデータベースに関する研究が、精力的に行われている。本論文では、曖昧なデータを取り扱うデータベースの一構築法として、曖昧さに関する知識を配列で表現する方法を提案し、それを SUSAN と名づけ実装した。SUSAN は単純な構成でありながら、複雑な有機的つながりを持ったデータ関係を柔軟に表現できる利点を持っている。我々は、実際にアパート紹介のモデルに SUSAN を応用し、その正当性も実証した。

1. はじめに

現在、人工知能的な手法を用いて計算機に推論を行わせる知識¹⁾²⁾ ベースをはじめとして、高機能なデータベースに関する研究が精力的に行われている。

我々は現在、曖昧なデータを取り扱うデータベースの検討を行っている。曖昧なデータとは、特定のデータに対して何らかの関連を持つデータの全集合としても捉えられる。関連が弱くなると、その曖昧さは増大すると考えられる。相互関連を有するデータを扱えるデータベースは、意味的つながりを持ったデータを処理できるという意味において、セマンティック・データベースの一形態であろう。以後、本論文ではこれを単にセマンティック・データベースとよぶことにする。

セマンティック・データベースの利点は、関連事項を含めてデータベースの検索を行う場合、自動的に関連物を検索するために、検索項目を並列に列挙する必要がなく、また、調べようとする項目について知識を持たないユーザも適切な調査が可能である点にある。

すでに我々は、セマンティック・データベースの一構築法として、誤り訂正符号を利用した ARES³⁾⁴⁾ と呼ばれるシステムの検討を行ったが、データのコード化に問題があることが判明した。

本論文では、この問題点を解決するためにデータベースの記憶データと別に曖昧さに関する知識を保持する新たな方法について論じ、具体的な実験を行い、有用性・実現性を示すことを目的としている。

まず、典型的なセマンティック・データベースの動作を示すとともに広い応用分野を紹介し、セマンティッ

ク・データベースに必要な曖昧さに関する知識を配列 (SA: Semantic Array) で表現することを提案する。次に、この SA を用いたセマンティック・データベース (SUSAN) の特徴を述べ、実際にアパート紹介のモデルに応用した例を示す。最後に、SUSAN とほかのシステムの比較を行い、今後の研究課題について述べる。

2. セマンティック・データベースの適用分野

たとえば、緑色のような色をした物を探す場合に、一般のデータベースでは、「緑色あるいは深緑あるいは黄緑の色をした物は何か?」というように、検索する項目を列挙しなければならない。それに対し、セマンティック・データベースでは、「緑色のような色をした物は何か?」というだけで、計算機が自動的に検索項目を決定し、検索を行う。

このように、セマンティック・データベースの特徴は、類似物を一緒に検索する点にある。この性質は非常に多くの分野で望まれている。

- 文献検索
- 不動産紹介
- 結婚相談
- 各種教育 etc.

たとえば文献検索において、「マイクロプログラム」の論文のサーベイを行う時は、「マイクロプログラム」に関連する項目を扱った論文も探したい。このような場合に、セマンティック・データベースは有効である。特に専門知識を必要とする分野で、セマンティック・データベースはその効力を最大限に発揮することができる。

3. セマンティック・データベースの原理

セマンティック・データベースでは、ある質問に対

[†] An Approach to the Database Processing Fuzzy Data by SHUJI SUGIHARA, JUN MIYAZAKI and HIDEO AISO (Department of Electrical Engineering, Faculty of Engineering, Keio University).

^{††} 慶応義塾大学工学部電気工学科

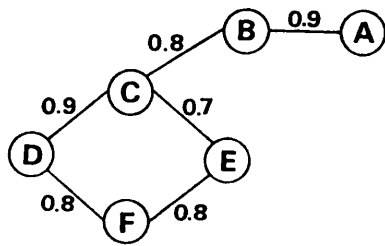


図1 セマンティック・ネットワークの例
Fig. 1 An example of Semantic Network.

し、その要求に合致する情報は勿論、それに関連する情報も応答する必要がある。そのためには、どの情報とどの情報が関連があるかを知識として持っておく必要がある。そして、その知識を用いて検索を行う。

3.1 知識の表現

一般にデータベースで扱う情報は多岐に及び、その情報間の関係もさまざまである。重さという属性における属性値間の関係は直線的であるし、色ではループ、家系では木構造になるであろう。また、地理データなどにおける位置関係は、単なる法則を持たず、複雑なデータ関係を持つことが予想される。これらさまざまな関係をうまく表現する方法として、グラフ表現がある。そして、知識をグラフとして表現したものをセマンティック・ネットワークと呼ぶことにする。

この場合、ノードはデータ項目を表わし、アークはデータ項目間に関係があることを示す。たとえば、図1ではデータAとデータBが関係のあることなどが示されている。

一般にデータベースのキーは複数の属性を持つ。各属性にそれぞれセマンティック・ネットワークを割り当てることにより、キー間の関係は、属性間の関係を含んだ有機的なつながりとなる。

3.2 セマンティック・データベースの動作

たとえば、図1のような関係でAに似た物を求める場合、出力はAだけでなくBもなければならない。そのために、図1のセマンティック・ネットワークでAを起点とし、アークをたどりBに目をつけ、次にデー

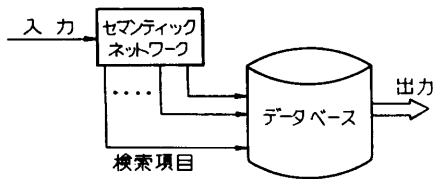


図2 提案するセマンティック・データベースの動作
Fig. 2 Proposed Semantic Database.

	A	B	C	D	E	F
A	1.0	0.9	0.7	0.6	0.5	0.4
B	0.9	1.0	0.8	0.7	0.5	0.5
C	0.7	0.8	1.0	0.9	0.7	0.7
D	0.6	0.7	0.9	1.0	0.6	0.8
E	0.5	0.5	0.7	0.6	1.0	0.8
F	0.4	0.5	0.7	0.8	0.8	1.0

図3 セマンティック・アレイの例
Fig. 3 An example of Semantic Array.

タベースをAとBの両方で検索を行う。(図2)

このような方法を用いれば、アークをどのくらいたどるかによって、曖昧さの程度を制御することができる。ただし、この方法ではデータベースの検索を行うたびにアークをたどるために、速度が遅くなる欠点を持つ。

そこで、実行時前にすべてのデータ項目間の関係をセマンティック・ネットワークから求めておき、これをデータベース検索時に用いる方法を考えた。この総当たりのデータ項目間の関係は、図3のように配列で表現される。これをSA (Semantic Array)と呼び、これを用いた本セマンティック・データベース・システムをSUSAN (Semantic database Using a Semantic Array Network)と名づけた。

4. SUSAN

4.1 SUSAN の特徴

SUSAN の特徴をまとめると次のようになる。

(1) 知識をデータと別に配列で持つこと

SA を用いることにより、柔軟性を失うことなくしに構造を単純化し、高速性を保つことができる。

(2) 関係値を正規化していること

データ項目間の関係の度合いを関係値と呼び、0~1の間の値を持たせる。このように関係値を可変量とすることは、セマンティック・ネットワークのアークを必ず一定量に見積もる場合に比べて、柔軟性を増し、論理長と物理長の対応がとれ、場所によって物理長の重みが異なるといった矛盾を解決することができる。

以降、データ項目 x_1, x_2 間の関係値を $R(x_1, x_2)$ と表記する。

$$0 \leq \forall x_1, \forall x_2 R(x_1, x_2) \leq 1 \quad \dots (1式)$$

ただし、関係値0は全く関係のないことを示し、関係値1は最も関係がある(自分自身との関係値 $R(x_1, x_1)$ である)ことを示す。

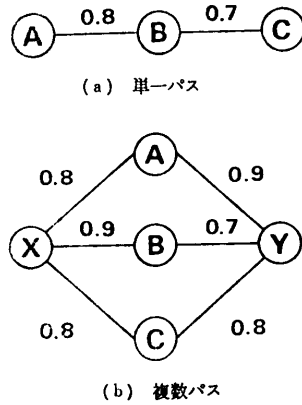


図 4 推論規則の例

Fig. 4 Examples of reasoning rule.

(3) 関係値に関し、推論規則を持つこと

セマンティック・ネットワークを SA に写像するために、セマンティック・ネットワークで直接表現されていない関係の関係値を推論により求める必要がある。この場合、推論された値が常に (1式) を満足しなければならない。SUSAN においては、次のような推論規則を用いる。

a. 単一パスの場合

データ項目 x_1, y 間および y, x_2 間の関係値 $R(x_1, y), R(y, x_2)$ が判明しているとき、 x_1, x_2 間の関係値 $R(x_1, x_2)$ は次式で求められる。

$$R(x_1, x_2) = R(x_1, y) \times R(y, x_2) \quad \dots\dots (2式)$$

たとえば 図 4(a) では、

$$R(A, C) = R(A, B) \times R(B, C) = 0.8 \times 0.7 = 0.56$$

単一パスで途中にノードが複数ある場合は、(2式) を再帰的に用いることにより推論される。

b. 複数パスの場合

求める x_1, x_2 間のあいだに複数のパスが存在するとき、関係値 $R(x_1, x_2)$ は次式で求められる。

$$R(x_1, x_2) = \text{Max}_{i=1, n} \{R(x_1, y_i) \times R(y_i, x_2)\} \quad \dots\dots (3式)$$

たとえば 図 4(b) では

$$\begin{aligned} R(X, Y) &= \text{Max} \{R(X, A) \times R(A, Y), R(X, B) \times R(B, Y), \\ &\quad R(X, C) \times R(C, Y)\} \\ &= \text{Max} \{0.72, 0.63, 0.64\} \\ &= 0.72 \end{aligned}$$

(4) 方向性を持った関係値とすること

たとえば、値段の安いことが明らかに望まれている分野において、二つの品物 x_1, x_2 が存在し、それら

に次式の関係が成り立つとする。

$$\text{price}(x_1) > \text{price}(x_2)$$

この場合に、 x_1 を望んだ時に x_2 が一緒に選ばれることは都合が良いが、 x_2 を望んだ時に x_1 も一緒に選ばれることは都合が悪い。このことは、価格という属性において、 x_1 から x_2 への関係値 $R(x_1, x_2)$ と x_2 から x_1 への関係値 $R(x_2, x_1)$ が異なることを示している。

逆に言えば、 $R(x_1, x_2)$ と $R(x_2, x_1)$ が異なることを認めれば、属性の持つ意味を SA に含ませることができることを示している。

$$R(x_1, x_2) \neq R(x_2, x_1) \quad \dots\dots (4式)$$

(5) 曖昧さの制御としての 2 種類の方法を併用していること

a. δ 制御

たとえば、図 3 の SA で、A と 0.9 以上の関係のあるものは A, B の 2 つであるが、これを 0.7 以上にすれば A, B, C の 3 つになる。したがって、この関係値の閾値 R を指示することによって曖昧さを制御することができる。これを δ 制御と呼ぶ。 δ を次式で定義する。

$$\delta = 1 - R \quad \dots\dots (5式)$$

b. θ 制御

一般にデータベースのキーは複数の属性より構成される。ユーザが S 個の属性の属性値を指定した時に、 θ 個の属性値を守り、 $\bar{\theta}$ 個の属性値を無視して検索を行ったとすると、 $\bar{\theta}$ の大きさによって出力の個数が変化する。これを θ 制御と呼ぶ。

$$\bar{\theta} = S - \theta \quad (\theta \leq S) \quad \dots\dots (6式)$$

SUSAN では、ユーザの指定により δ 制御が行われ、出力が 0 の場合に自動的に θ 制御が行われる。

(6) 連想結果の評価と指定個数出力

ある曖昧さの許容内に入る情報は、理想の情報にどれだけ近いかを評価され、ソートされて出力される。したがって、要求に近いものから順に出力される。

また、SUSAN のユーザは δ 方向の曖昧さのほかにも出力数を指定することができる。したがって、ユーザは余分な情報を省いて必要な情報のみを受け取ることができる。

4.2 SUSAN の拡張機能

SUSAN の特徴は以上のようなものであるが、拡張によって、次のような機能を持つことが可能である。

(1) SA の自動学習機能

図 5 のように、ユーザに対し御意見申受箱を用意

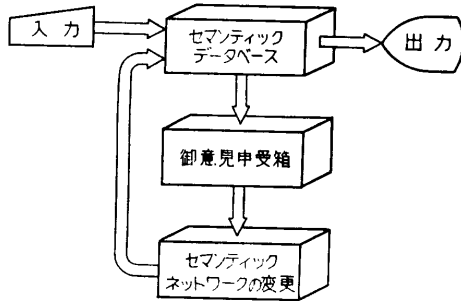


図 5 自動学習機能

Fig. 5 Automatic learning function.

し、セマンティック性に関する苦情を集める。この値がある閾値を越したら、それを是正するように SA を自動的に改変する機構を設ける。この機構により、真にユーザの抱えているデータのセマンティック性をセマンティック・データベースに反映することが可能となる。

(2) セマンティック・データベースの個人適用化

同じデータに対しても、事務屋の考える意味と技術屋の考える意味は異なる。したがって、SA を事務屋用と技術屋用に別々に用意すれば、同じデータベースがあたかも別々に存在するようになる。これを個人にまで押し広げれば、統合データベースもそのユーザにとっては、自分自身のデータベースのように見ることが出来る。

また、この機構は SA の保護機構としても働く。つまり、先に述べた自動学習機能を悪用し SA を破壊しようとする者に対し別に SA を割り当てれば、元の SA は破壊されない。

5. SUSAN の実験システム

前章の考えに従い、SUSAN の実験システムを作成した。その目的は、知識を SA で表現した SUSAN

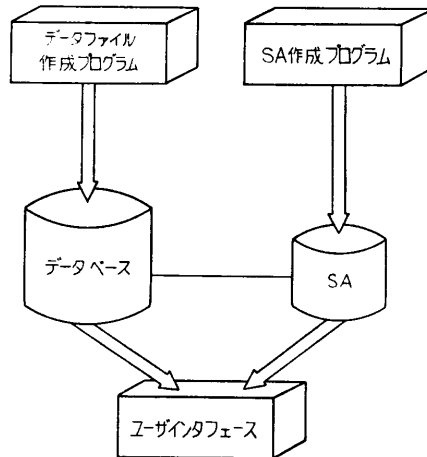


図 6 Susan の構成

Fig. 6 Configuration of SUSAN.

が有用であり、かつ正しいセマンティック性を持って動作することを示すことにある。

対象モデルはアパート・マンション紹介とした。その理由は、アパート紹介においては、マンマシン・インタフェースにおける曖昧性がシステムの性能に大きな影響を与えうること、またデータ項目間の関係がほぼ明解であり実験として適していることなどである。

データ・モデルは自由なアクセス性・並列処理の可能性からリレーショナル・モデルとした。

5.1 実験システムの構成と動作

実験システムは図 6 のように 3つのプログラムより構成され、以下のような動作を担当している。

(1) データファイル作成プログラム

アパート・マンションに関するデータベースのデータファイルの作成を目的とする。レコードの構成は図 7 に示すようなものであり、2つのリレーションに正規化している。Cリレーションはアパートの棟全

Cリレーション

登録番号	名称	電話番号	場所	距離	性別	学生	結婚

Rリレーション

レコード番号	登録番号	室番号	家賃	タイプ	大きさ	トイレ	バス	契約

図 7 レコードの構成

Fig. 7 Record configuration.

体に共通する情報を、Rリレーションでは室ごとに異なる情報を扱う。

各属性値については、次に述べるSA作成プログラムで作成したコード化表に従ってコード化を行い、データの正当性確認も行っている。

(2) SA 作成プログラム

図1に示されるような知識を、図3に示されるようなSAに変換することを目的とする。具体的には、まずシステム作成者が、セマンティック・ネットワークに対応させて、2つのデータ項目とそれらの関係値を入力する。すると、SA 作成プログラ

```

** SHIBUYA
(0.9) EBISU HARAJUKU
(0.8) YOYOGI MEGURO
(0.7) SHINJUKU GOTANDA
(0.6) OOSAKI TAKADANOBABA SHINOOKUBO
(0.5) SHINAGAWA MEJIRO
(0.4) TAMACHI HAMAMATSUCHOU IKEBUKURO
(0.3) SUGAMO
(0.2) TABATA NIPPOLI
(0.1) UGUISUDANI UENO
    
```

図 8 SA 作成プログラム出力表示例

Fig. 8 An example of result by SA generator program.

```

HELLO, I WILL HELP YOU TO SEARCH YOUR APARTMENT.
① ARE YOU MALE OR FEMALE ? M OR F : M:
ARE YOU A STUDENT ? Y OR N : Y:
ARE YOU MARRIED ? Y OR N : N:
HOW MUCH HOUSE-RENT PER MONTH ? : 80000:
HOW STRONGLY MUST I RECEIVE YOUR REQUEST ? ← ③
1.ABSOLUTELY 2.STRONGLY 3.NOT SO STRONGLY 4.WEAKLY :2:
IN WHICH NEIGHBORING STATION DO YOU WANT TO LIVE ? :YOYOGI:
HOW STRONGLY MUST I RECEIVE YOUR REQUEST ?
1.ABSOLUTELY 2.STRONGLY 3.NOT SO STRONGLY 4.WEAKLY :2:
② HOW FAR IS IT FROM THE STATION ON FOOT ? :10:
HOW STRONGLY MUST I RECEIVE YOUR REQUEST ?
1.ABSOLUTELY 2.STRONGLY 3.NOT SO STRONGLY 4.WEAKLY :3:
WHAT ROOM TYPE DO YOU WANT ? EX. 2DK 1ROOM :2DK:
HOW STRONGLY MUST I RECEIVE YOUR REQUEST ?
1.ABSOLUTELY 2.STRONGLY 3.NOT SO STRONGLY 4.WEAKLY :3:
HOW LARGE ARE THE ROOMS ? : ANY: ← ④
⑤ DO NEED A PRIVATE TOILET BY ALL MEANS ? Y OR N :Y:
DO NEED A PRIVATE BATH BY ALL MEANS ? Y OR N :Y:
⑥ * THE OBJECT THAT MEET YOUR REQUESTS ARE 10 APARTMENTS.
HOW MANY APARTMENTS DO YOU WANT TO KNOW ? :3:
=====
** COPO-MATSUMOTO ROOM NO.=520
PLACE=SHINJUKU DISTANCE=10 MIN.
MONTH FEE=85000 ROOM TYPE=2DK
ROOM SIZE=6 8 TOILET BATH
FOR ANYBODY TEL=109-8765
=====
    
```

図 9-1 アパート紹介応用例

Fig. 9-1 Demonstration in apartment consulting.

ムがこのデータを基に、(2式)および(3式)の推論規則を用いてすべてのデータ項目間の関係を推論し、SAを作成する。

本アプリケーションにおいてSAが必要であるのは、図7におけるCリレーションの場所・距離、Rリレーションの家賃・タイプ・大きさの5属性である。それら各属性に対し、このプログラムを用いて、属性値間の関係を示すSAをそれぞれ作成した。ただし、属性値は小数第2位以下を切り捨てた。

また、このプログラムは、SAを図8のように見やすい形で表示することができ、もしも推論結果が人間の感覚と異なる時はTSSにより部分修正ができる

など、さまざまな機能を有している。

(3) ユーザ・インタフェース

ここで対象とするユーザは初心者であり、データベースに関し内容を知らないと仮定している。そのため、ユーザはTSS端末を用い計算機の質問に答えるだけで希望のアパートを紹介できるようにした。その動作例を図9に示す。下線を引いた部分がユーザの応答であり、残りは計算機の応答である。ただし、番号は以下の説明文に対応している。

① 性別・学生・既婚などを尋ねている。これはアパートの大家さんの要求であり、これに対し曖昧な処理は行わない。

```

=====
.....
.....
=====
ARE YOU SATISFIED THESE INFORMATIONS ? Y OR N :N:
WHICH REQUEST WILL YOU CHANGE ?
  1. MONTH FEE  2.PLACE  3.DISTANCE  4.ROOM TYPE
  5.ROOM SIZE  6.TOILET ROOM  7.BATH ROOM  :2:
⑧ IN WHICH NEIGHBORING STATION DO YOU WANT TO LIVE ? :UENO:
HOW STRONGLY MUST I RECEIVE YOUR REQUEST ?
  1.ABSOLUTELY  2.STRONGLY  3.NOT SO STRONGLY  4.WEAKLY  :1:
ARE THERE ANY CHANGES ? Y OR N :N:
⑨ * THE OBJECT THAT MEET YOUR REQUESTS ARE NOT FOUND.
  BUT I RECOMMEND YOU THESE 23 APARTMENTS.
HOW MANY APARTMENTS DO YOU WANT TO KNOW ? :2:
=====
** UDAGAWA-SOU          ROOM NO.=38
  PLACE=UGUISUDANI     DISTANCE=10 MIN.
  MONTH FEE=85000      ROOM TYPE=3DK
  ROOM SIZE=4.5 6 8    TOILET BATH
  FOR ANYBODY          TEL=931-5678
=====
.....
.....
=====
ARE YOU SATISFIED THESE INFORMATIONS ? Y OR N :Y:
I AM GLAD, IF I CAN HELP YOU.
THANK YOU. GOOD-BY.

```

図 9-2 アパート紹介応用例

Fig. 9-2 Demonstration in apartment consulting.

② 家賃・場所・距離・室タイプ・大きさに対する質問であり、これらについて曖昧処理を行う。

③ ユーザの要求の強さを尋ねている。これにより関係値の閾値を決め、曖昧さに対する θ 制御を行う。具体的には、現在のシステムでは閾値が、absolutely が 1.0, strongly が 0.8, not so strongly が 0.6, weakly が 0.4 に設定されている。

④ 特定の希望はなく、何でも良いことを示す。

⑤ トイレ・風呂に対する質問である。応答がYのときはあるものが検索されるが、Nのときはあるものもないものも検索される。

⑥ θ 制御により 10 軒見つかったことを示している。

⑦ ユーザの指定した軒数分だけ、要求に近いアパートから順に紹介している。

⑧ ユーザに要求の変更があるかを尋ねている。

⑨ 曖昧さに対する θ 制御だけでは出力が0であったので、 θ 制御を併用したことを示している。その結果、23軒のアパートが見つかった。現在、 θ 制御においては $\theta=1$ までしか行わず、その場合にも出力が0の時は、But 以下の出力は行わない。

5.2 実験システムの評価

アパートデータとして、棟数約 30 軒、室数約 100 室を用意し、質問に対する応答として擬似乱数を用いた場合に、一般のデータベースと比較して、実験システムがどのような性能を持つかを実験値を用いて示す。ただし、ここで比較する一般のデータベースとは、単に SUSAN の曖昧性をなくしたものであり、何々以上何々以下といった範囲を指定することはできないものと仮定した。

(1) 曖昧さに対する出力率

図 10 に、1回の質問でデータベースが答える室数が曖昧さに対してどのように変化するかを示す。ここで出力率とは(応答室数)/(記憶室数)である。このグラフを見てわかることは、当然ではあるが、曖昧さを増すと出力数が増すことである。また、一般のデータベース($\delta=0.0$)では出力率は0に近く、一回の質問ではほとんど出力されないことが判る。

(2) 速度

図 11 に希望数の出力を得るのに要するアクセス時間が曖昧さに対してどのように変化するかを示す。このグラフを見てわかることは、曖昧さが増すとアクセス時間が減るが、 δ が0.3以上になるとアクセス時間はほぼ一定することである。これは、ある程度以上曖昧

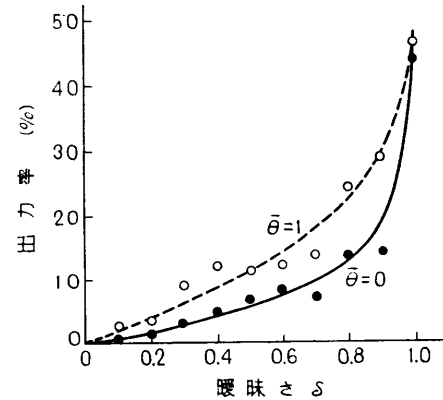


図 10 曖昧さに対する出力率

Fig. 10 The no. of answered rooms/Total no. of rooms VS. Ambiguity.

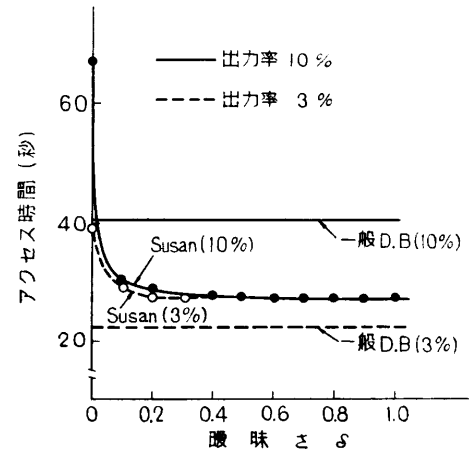


図 11 曖昧さに対するアクセス時間

Fig. 11 Access time VS. Ambiguity.

さが増すと、一回の質問で希望の出力数以上の出力が得られることを示している。さらに、一般のデータベースではアクセス時間は希望の出力数に大きく左右されるのに対し、SUSANではほぼ一定していることが判る。つまり、少ない出力数を得るには一般のデータベースの方が速いが、多くの出力数を得るにはSUSANの方が速い。この実験システムの場合、臨界点は出力率約5%であった。

(3) 容量

SUSANの場合は一般のデータベースと違って、データベースのほかにSAを持っていないからならぬ。アパート1軒に対し平均8室あると仮定して、この割合を示したのが図 12である。これによれば、100室以下では約1/4をSAが占めるが、500室程度で5%、それ以上の室数ではさらに占有率が下がる。これ

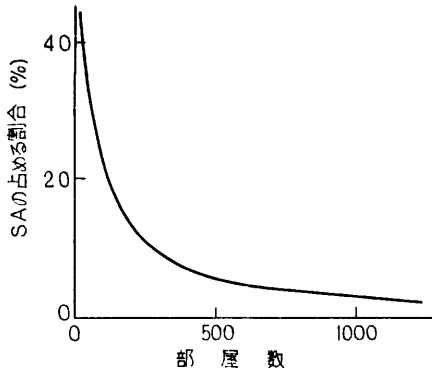


図 12 全データ量に対する SA の占める割合
Fig. 12 Percentage of SA in the whole data.

により、SA の容量上のオーバーヘッドは、それほど大きくないことが判る。

このように SUSAN は、それほど多くの容量増大なしに、速度も一般のデータベースとほぼかわりない状態で、曖昧性の機能を発揮することができる。

6. ほかのシステムとの比較

ARES や Fuzzy データベース^{6),7)}といったデータベースは、SUSAN と同様に曖昧性を持った推論を行うことができる。ARES の特徴は、データ間の関係をコード内に写像し、誤り訂正により特徴把握を行う点にあり、有機的なデータ間の関係を表現できる利点がある。また、Fuzzy データベースの特徴は、Fuzzy 理論を基礎としている点であり、曖昧さの表現の柔軟性に利点を持っている。

一方、SUSAN は、各属性に SA を割り当てることにより有機的なデータ関係を表現しており、また、曖昧さの表現は幾分かこちないが、属性値を等級分けすることなく、属性値間の複雑な関係を柔軟に表現できるという利点を持っている。さらに、SUSAN は既存のデータベースに SA を付加することにより、容易にセマンティック・データベースを構築することができるという点において、ほかのシステムにない高い実現性を持っている。

7. 今後の研究課題

SUSAN には、次に述べるような多くの本質的課題も残されている。

(1) SA の容量は、対象とする属性値の数の 2 乗に比例する。したがって、属性値の数が多すぎると、システムは現実的でなくなる。これを避けるために

は、図書分類法のような階層構造を持った分類が必要となる。その場合、上部階層と下部階層のリンクはどのようにとり、どのような推論を行うべきか。

(2) 現時点では推論は同一属性内においてのみ行われたが、属性間にまたがる推論はできないのか。これにより、記憶データの圧縮はできないのか。

(3) SUSAN はデータ検索を行うユーザに対してサービスを強化することができたが、データベース作成者には何の便宜も計っていない。データ入力を手助けするような実用的方策はないのか。

(4) 今回のシステムは本大学 FACOM M-180 を使い、PL/I で実験を行ったが、実際の DBMS を用いた場合は、下層スキーマに対しどのような影響をもたらすのであろうか。

(5) SUSAN の動作は非常に単純である。したがって、各 VLSI は単純な同一作業を行うが、全体としてはパイプライン・並列処理を行い、高速に出力を出す VLSI セルラシステムが考えられるのではないだろうか。

8. む す び

曖昧なデータ処理を行うセマンティック・データベース SUSAN は、自動的に類似物を検索する能力を持っており、一般のデータベースに比べて使いやすく、ユーザに対するサービスの強化が可能である。SUSAN の特徴は、曖昧さに関する知識を配列 (SA: Semantic Array) で表現している点であり、非常に単純な構成となっている。

SA は、すべての属性値間の関係を総当たりで表現しているため、非常に複雑な関係にも対処でき、各属性性にこの SA を割り当てることにより、データ間の有機的な関係を表現することが可能である。さらに、SA を用いる利点は、既存のデータベースに SA を付加することにより、容易にセマンティック・データベースを構築できる点にある。

我々は、実際にアパート紹介のモデルに SUSAN を応用し実験を行った結果、まだ多くの研究課題が残されているものの、その有用性・実用性は高いという確信を得た。

なお、本研究にご協力頂いた本研究室の武藤佳恭さん、足立佳彦君をはじめとする皆様に感謝します。

参 考 文 献

- 1) Davis, Randall, et al.: Production Rules as a

- Representation for a Knowledge-Based Consultation program AI 8, pp. 15-45 (1977).
- 2) Weiss, S. M. et al.: A Model-Based Medical Decision-Making, AI 11, pp. 145-172 (1978).
 - 3) Feigenbaum, E. A. et al.: On Generality and Problem Solving: A Case Study Using the DENDRAL Program, Machine Intelligence Vol. 6.
 - 4) 市川忠男, 小西祐蔵: セマンティックな語検索にみる連想プロセッサ ARES の振舞い, 電子通信学会, AL 78-50.
 - 5) Ichikawa, Tadao, et al.: A Multimicroprocessor ARES with associative processing capability on semantic data bases, AFIP Vol. 47, pp. 1033-1039 (1978).
 - 6) 三重野博司, 戸内順一: 推論過程を Fuzzy 化した自然言語システム — 建築材料選定を対象として —, 情報処理学会論文誌 Vol. 21, No. 6, pp. 507-511 (Nov. 1980).
 - 7) 馬野元秀: Fuzzy 概念とデータベース, 数理学, No. 191 (May 1979).
(昭和 55 年 12 月 18 日受付)
(昭和 56 年 4 月 27 日採録)