

1H-3

あいまい要求のための 検索システム

野村 恭彦 阿倍 博信 小高 俊之 平岩 真一 横山 光男 松下 温
(慶應義塾大学 理工学部)

1. はじめに

近年、あいまいな要求に対するデータベースの検索、というものが重要視されてきている[1]。それは、人間の要求は、必ずしも厳密なものではないからである。何かを買おうとした時などがそのいい例で、その意志決定は非常にあいまいなものであろう。例えば5万円以下のCDプレーヤーを買おうとしているときでも、6万円で非常に良い品があれば、買う気になってしまうかもしれない。

また我々の要求は、数値的にはっきりしたものよりも、こういった感じのもの、といったような形容詞的に表現されるものが多い。例えば車を選ぶときに、大きい車がいいとか、室内がゆったりした車がいいといった要求は自然なものだが、全長4200mm以上の車がいいとか、室内長何m以上の車がいいとかいった要求は、かなり我々の感覚にあわないものになってしまうだろう。

このような要求に答えるために、データベース検索にファジイ関係を用いたシステムを作成したので、ここに報告する。

2. 検索システム

2.1 試作システム

我々は、試作システムとして自動車データベースを構築した。大きさ・経済性・パワー等の要求から、希望にあった自動車を検索してくるシステムである。システムの構築は、Mac II CxのSuperCard上で行った。

2.2 要求の入力

まず条件選択画面で、検索したいアトリビュートをクリックする事により選ぶ[2]。すると、要求入力画面に変わるので、要求の入力を3つのアイコンから1つを選ぶことによって行う。この3つのアイコンとは、/大きい・中位・小さい/や/上級者向・中級者向・初級者向/などのように、形容詞で表された条件である。要求を入力すると再び条件選択画面に戻るので、重ねて要求がある場合は、複合条件の入力を同様に行

う。そして、全ての条件を入力したなら、検索ボタンをクリックする。次に要求の重要度の入力を行う。これにより、複合条件の重みつきの検索が行われることになる。

2.3 アトリビュートの作成

我々の検索システムでは、データ内にファジイなものを用意せずに、クリスプなアトリビュートからファジイなアトリビュートを導出することにした。そのためにn次元座標を考え、座標軸に各アトリビュートをとり、そのn次元座標に全てのデータをプロットした。そしてn次元座標の $X_1 = X_2 = \dots = X_n$ という直線上に、平均演算子を用いて各データを落とした[3]。このとき、3つの平均演算子(算術平均、幾何平均、調和平均)を条件の種類によって使い分けた(図1)。これにより、データの更新時にファジイ関係を入力する必要がなくなった。

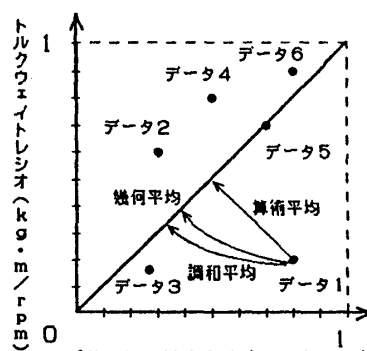


図1. $X_1 = X_2$ 直線への帰結方法

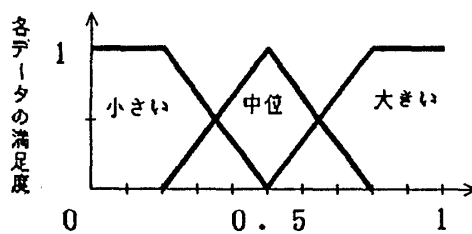


図2. メンバシップ関数のデフォルト

A Retrieving System for Ambiguous Requirements

Takahiko Nomura, Hironobu Abe, Toshiyuki Odaka, Shinichi Hiraiwa, Teruo Yokoyama, Yutaka Matsushita
Faculty of Science and Technology, Keio University

2.4 満足度の算出

要求の入力によって、そのアトリビュートに関する満足度が計算される(図2)。複合条件の場合、再びn次元座標を考え、各座標軸に、各アトリビュートの満足度をとる。そして、入力された重み付けにあわせて $X_1 = X_2 = \dots = X_n$ 直線上に、平均演算子のANDを用いて落として、そのデータの満足度を算出する。

2.5 平均演算子

平均演算子には、算術平均と幾何平均と調和平均の3つを用いた。一般式を次に示す。

算術平均

$$\frac{\sum_{i=1}^n X_i}{n}$$

幾何平均 (ANDの場合)*

$$\left(\prod_{i=1}^n X_i \right)^{1/n}$$

調和平均 (ANDの場合)*

$$n / \sum_{i=1}^n (1/X_i)$$

(*ORの場合は、その双対なものをとる。)

なぜ平均演算子を用いたかという点、ANDに単なる最小値をとってしまうと、(0.8,0.2)も(0.2,0.2)も、満足度は同じになってしまう。しかし実際は、最小値でない値も、満足度に全く関わりがないわけではないだろう。そこで、この違いを表現できて、なおかつ重み付けの容易であるために、これを用いることにしたのである。この平均演算子を使うことによって、下位の方の順位も決まり易くなった。

それぞれの平均演算子の使い分けだが、複数条件のどれもを満たしていなければならないような、厳しいAND条件の時は調和平均を使い、それより少し甘い条件の時は幾何平均を使う。ANDともORともつかない条件のときは、算術平均を用いることにした。

3. ブラウジング環境

3.1 結果表示画面

まず、検索結果の表示画面を図3に示す。現在注目している中心のデータだけが、細かい情報を与えている。上下を示すボタンを押すと、表がスクロールし、また、順位をクリックすることによって、そのデータを中心に持ってくる事ができる。それから、中心にきているデータの順位をクリックすると、そのデータについての、さらに詳しい情報が別ウィンドウに表示される。また、チェックボックスによって、データに印をつけることができる。図には含まれていないが、さ

らに2つのウィンドウが開いており、1つにはその車の写真が、もう1つには検索条件が表示されている。そして、メニューバーにいくつかの特別なブラウジング機能を用意した。

<input type="checkbox"/>	1	No.19	NISSAN FAIRLADY Z	¥3950000	↑
<input type="checkbox"/>	2	No.18	NISSAN CIMA Type2	¥4329000	
<input type="checkbox"/>	3	No.3	TOYOTA SUPRA 3000GT Turbo Limited Aerotop Spec : S6 DOHC Turbo cc : 2954 cc AT/MT: 5MT Length : 4620 mm Width : 1745 mm Height : 1300 mm Max Power : 240 ps/ 5600 rpm Max Torque : 35 kg m/ 3200 rpm lOmode : 8.3 km/l	¥3471000	
<input checked="" type="checkbox"/>	4	No.8	TOYOTA MR2 2000GT	¥2778000	↓
<input type="checkbox"/>	5	No.7	TOYOTA CELICA GT-FOUR	¥2685000	
<input type="checkbox"/>	6	No.20	NISSAN SKYLINE GTS-4	¥2790000	

図3. 検索結果の表示画面

3.2 ブラウジング機能

Sort...各アトリビュートごとに降順・昇順の並び換えを行う。

Go...最初や最後のデータやチェックしたデータへと、ジャンプできる。

Similar...現在中心にきているデータに、最も似ているものを3つ、別ウィンドウに表示する。

Check...チェックしたデータを抽出する。

Delete...中心にあるデータを消去する。

Setting...現在検索されているデータの中で、さらに2次検索を行う。

Define...メンバシップ関数の、ユーザーによる再定義を行う。
...etc.

4. おわりに

あいまいな要求に答えるための検索システムを提案した。当システムの最大の特徴は、データベースの知識のあまりないユーザーにも、あいまいな要求に対する検索が容易に行えることである。このシステムでは、自動車の写真は表示したにすぎないが、画像による検索まで進めば、マルチメディア環境として面白くなるのであろう。

5. 参考文献

- [1] Maria Zemankova, Abraham Kandel, "Implementing Imprecision in Information Systems", Information Science 37, 107-141 (1985).
- [2] 堀尾, 津田, 平川, 田中, 市川, "アイコンニックブラウザの開発", 情報処理学会データベースシステム研究会資料, 89-DBS-70(4), 1989.
- [3] 平岩, 阿倍, 小高, 野村, 横山, 松下, "拡張演算子のあいまい要求処理への適用", 本大会