第4章　ファジィマルチ集合の情報システムへの応用

4.1　ラフ集合への応用

知識に対する基本的性質あるいは概念というのは類別 (classification) とカテゴリである。知識に関する研究はもともと哲学の問題であった。また認識の問題を取り扱う認知科学の中心テーマである。しかし、この論文で議論しようとするのはある情報表現の手段でラフ集合の概念を利用して表現しようとするのである。ラフ集合の基本概念は類別と近似である。ラフ集合は目標とするクラスに従属する集合をある属性値を満たす集合でどのように近似できるかという問題から出発している。これは目標となるクラスに従属する標本で表される集合が概念を表す属性と属性値との対象を満たす集合でどのように表すことができるかという問題とみなせる。

Pawlak によるラフ集合では，全体集合 X の分割 $E_1, \ldots, E_s$:

$$\bigcup_{i=1}^{s} E_i = X, \quad E_i \cap E_j = \emptyset \ (i \neq j)$$

が与えられたと仮定し，集合の上近似 R^* と下近似 R_* について考察する [22, 27, 21]. また，この分割によって生じる同値関係 R によって $x \in X$ が属するクラスを $[x]_R$ と書く．すなわち，$x \in E_i$ となる E_i について $[x]_R = E_i$ である．また，

$$X/R = \{E_1, \ldots, E_s\}$$

と表す．

Dubois と Prade [5] はファジィ集合 D に対してラフファジィ集合として上近似

$R^*(D)$ および下近似 $R_*(D)$ を次のように定義している．ただし，ここで扱うのは有限集合であるので，元の定義 [5] における sup と inf を max と min で置き換えている．

$$\mu_{R^*(D)}(x) = \max_{y\in[x]_R} \mu_D(y) \quad (4.1)$$

$$\mu_{R_*(D)}(x) = \min_{y\in[x]_R} \mu_D(y) \quad (4.2)$$

クリスプ集合であるかファジィ集合であるかを問わず，上近似と下近似は同値類 $E_1, ..., E_s$ を指定すれば充分であることに注意しよう．次の例を参照されたい．

例 4.1 $X = \{x_1, ..., x_8\}$ における同値関係 R が分割 $E_1 = \{x_1, x_2, x_3\}$, $E_2 = \{x_4, x_5\}$, $E_3 = \{x_6, x_7\}$, $E_4 = \{x_8\}$ で与えられるものとする．クリスプ集合 K とファジィ集合 D がそれぞれ

$$\begin{aligned} K &= \{x_1, x_2, x_6, x_7\}, \\ D &= \{(x_1, 1), (x_2, 0.5), (x_6, 0.8), (x_7, 0.9)\} \end{aligned}$$

で与えられるものとする．

$$\begin{aligned} R^*(K) &= \{x_1, x_2, x_3, x_6, x_7\} \\ R_*(K) &= \{x_6, x_7\} \\ R^*(D) &= \{(x_1, 1), (x_2, 1), (x_3, 1), (x_6, 0.9), (x_7, 0.9)\} \\ R_*(D) &= \{(x_6, 0.8), (x_7, 0.8)\} \end{aligned}$$

であるが，これらを

$$R^*(K) = \{E_1, E_3\} \quad (4.3)$$

$$R_*(K) = \{E_3\} \quad (4.4)$$

$$R^*(D) = \{(E_1, 1), (E_3, 0.9)\} \quad (4.5)$$

$$R_*(D) = \{(E_3, 0.8)\} \quad (4.6)$$

と表しても意味としては同じである．厳密には，後の表現では前の記号 R^*, R_* とは違った記号を用いるべきであるが，混乱する恐れがないので，同じ記号ですますことにする．

関数 f を X から同値類 X/R の上への自然な写像

$$f(x) = [x]_R \tag{4.7}$$

とする．例 4.1 における後の表現のように E_i を用いた場合，上近似はクリスプ集合 K とファジィ集合 D について

$$\begin{aligned} R^*(K) &= f(K) \\ R^*(D) &= f(D) \end{aligned}$$

と表される．ここで，右辺は通常の集合を表す．これに対して下近似は補集合 K^C, D^C を用いることによって

$$\begin{aligned} R_*(K) &= f(K^C)^C \\ R_*(D) &= f(D^C)^C \end{aligned}$$

と表される．

これらの考察から，X のマルチ集合とファジィマルチ集合について $f[\cdot]$ を適用してみることが考えられる．そこで，X のクリスプマルチ集合を L, ファジィマルチ集合を A とし，(4.7) で定義された自然写像によって

$$R^*[L] = f[L] \tag{4.8}$$

$$R^*[A] = f[A] \tag{4.9}$$

と定義しよう．(4.8), (4.9) における右辺はマルチ集合を表している．

注意. 命題 3.4 に示したように，ファジィマルチ集合の像はクリスプマルチ集合の像の一般化であるので，(4.9) は (4.8) を一般化している．従って，上の定義は (4.9)

だけで充分である．よって以下では，ファジィマルチ集合についてのみ結果を述べる．クリスプマルチ集合については，メンバーシップが0または1に限られたファジィマルチ集合と考えればよい．

命題 4.1. R^*, R_* の表現として，$E_1, ..., E_s$ を用いるものとする．A を X のファジィ集合とするとき，

$$
\begin{aligned}
R^*(A) &= \mathcal{P}(R^*[A]) && (4.10)\\
R_*(A) &= \mathcal{P}(R^*[A^C])^C && (4.11)
\end{aligned}
$$

が成立する．

例 4.2 例 4.1 のファジィ集合 D について

$$
\begin{aligned}
R^*[D] &= \{(E_1,1),(E_1,0.5),(E_3,0.8),(E_3,0.9)\}\\
&= \{\{1,0.5\}/E_1,\{0.9,0.8\}/E_3\}
\end{aligned}
$$

より，

$$\mathcal{P}(R^*[D]) = \{(E_1,1),(E_3,0.9)\}.$$

となり例 4.1 の (4.5) に一致する．また，

$$
\begin{aligned}
D^C = \{&(x_2,0.5),(x_3,1),(x_4,1),(x_5,1),\\
&(x_6,0.2),(x_7,0.1),(x_8,1)\}
\end{aligned}
$$

から，

$$
\begin{aligned}
R^*[D^C] &= \{(E_1,0.5),(E_1,1),(E_2,1),(E_2,1),\\
&\quad (E_3,0.2),(E_3,0.1),(E_4,1)\}\\
&= \{\{1,0.5\}/E_1,\{1,1\}/E_2,\{0.2,0.1\}/E_3,\\
&\quad \{1\}/E_4\}.
\end{aligned}
$$

よって,

$$\begin{aligned}\mathcal{P}(R^*[D^C]) &= \{(E_1,1),(E_2,1),(E_3,0.2),(E_4,1)\},\\ \mathcal{P}(R^*[D^C])^C &= \{(E_3,0.8)\}\end{aligned}$$

となり, 例 4.1 の (4.6) に一致する.

命題 4.1 の証明

上近似に関する第 1 の式は, A が通常のファジィ集合のとき $\mathcal{P}(A) = A$ であることに注意して命題3.5 の第 1 式を適用すればよい. 第 2 式を示すため, グレード列による表現において

$$I(x) = \max\{i : \mu_A^i(x) \neq 0\}$$

とおく. ただし, $\mu_A^1(x) = 0$ ならば $I(x) = 0$ とする. また,

$$R^*[A] = \{(\mu^1(E_i), ..., \mu^{I(E_i)}(E_i))/E_i\}_{i=1,..,s}$$

とする. 全体集合 X が有限なので, $I(E_i) < \infty$ である. A^C の定義から

$$R^*[A^C] = \{(1 - \mu^{I(E_i)}(E_i), ..., 1 - \mu^1(E_i))/E_i\}_{i=1,..,s}$$

この式は $I(E_i) = 0$ のときも成立していることに注意しよう. よって

$$\mathcal{P}(R^*[A^C]) = \{1 - \mu^{I(E_i)}(E_i)/E_i\}_{i=1,..,s}$$

すなわち

$$\mathcal{P}(R^*[A^C])^C = \{\mu^{I(E_i)}(E_i)/E_i\}_{i=1,..,s}.$$

一方, $\mu^{I(E_i)}(E_i)$ は E_i に写される x のうち最小のメンバーシップ値を表しているから

$$\mu_{R_*(A)}(E_i) = \mu^{I(E_i)}(E_i)$$

である. よって証明された.

命題 4.1 の証明にも示されているが，$R^*[A]$ から下近似 $R_*(A)$ を計算するためには，補集合演算を行うまでもなく，$R^*[A]$ におけるグレード列の最小値をとればよい．このように，$R^*[A]$ は上近似と下近似の両方の情報を含んでいる．また，最初に述べたように，$R^*[A]$ は単純な逐次演算で得られるため，通常の上近似と下近似を計算する前段階の情報表現を表していると考えられる．

4.2 ファジイ関係データベース

ファジイ関係データベースシステム [16, 29, 33] は、古典的な関係データベースシステム [28] の機能を拡張された分野であるから伝統的なデータベースが応用されるすべての場面で使用できる。現実の世界に存在する曖昧なデータとか大量のデータを効率よく蓄積、処理、検索するためには従来のデータモデルでは足りない点があるのでファジイデータベースが提案されている。人間が直接関係する分野にはそのデータを明確に定義することができないし、定義する必要もないあいまいなデータが多数存在する。そしてデータにはあいまいさはなくてもあいまいな質問をする場面も多い。そこで曖昧なデータと質問を可能性分布 (確率分布) やファジイ集合により表現することができる。本論文で提案しようとするのはファジイ関係データベースの応用としてファジイマルチ集合を利用した SQL の演算 [3, 28] に対する応用としてファジイマルチ集合に関する 2 種類の像のうち $f[A]$ に対する関係について考えられる。

4.3 SQLへの応用

関係データベースは集合論にもとづいているが，その問い合わせ言語SQLはマルチ集合を扱っている [3]．このことは，$f[A]$ に対応する操作がSQLに含まれていることを意味している．そこで，ファジィデータベースにおける問い合わせと $f(A)$, $f[A]$ との関連に言及しておこう．

ここでは，関係を $K \subseteq X = X_1 \times \cdots \times X_h$ で表す．当面 K はクリスプで，かつマルチ集合ではないとする．

マルチ集合が生じる典型的な例は，射影 $\pi_I(K)$ [28] である．ここで，I は $\{1, 2, ..., h\}$ の部分列 $i_1, ..., i_\ell$ を示すものとする．すなわち，$x = (x_1, ..., x_h) \in K$ に対して，

$$\pi_I(x) = (x_{i_1}, ..., x_{i_\ell})$$

である．このとき，$f(x) = \pi_I(x)$ とおけば，$f[K]$ と $f(K) = \mathcal{P}(f[K])$ の両方が考えられる．

問い合わせ言語SQLでは $f[K]$ に対応するのは SELECT ALL であり，

SELECT ALL $i_1, ..., i_\ell$ FROM K

と書ける．$f(K)$ に対応するのは SELECT DISTINCT であり，

SELECT DISTINCT $i_1, ..., i_\ell$ FROM K

と書かれる．

次に，K はマルチ集合と仮定しよう．$\mathcal{P}(K)$ は

SELECT DISTINCT $*$ FROM K

と書ける．また，$f[K]$ に対応するのはやはり SELECT ALL である．ところが，SELECT DISTINCT は結果が通常の集合であることを要求しているので，$f(K)$ ではなく，$\mathcal{P}(f(K))$ あるいは $\mathcal{P}(f[K])$ に対応している．

例 4.3 関係スキーマ $\mathbf{K}$(C,Y) を考える．タプルの集合を

$$K = \{(C_1, a), (C_2, a), (C_1, b), (C_3, a), (C_2, b), (C_3, b)\}$$

とし，属性名が C,Y であるとする．

SELECT ALL Y FROM **K**

の結果はマルチ集合 $\{a, a, b, a, b, b\}$ であり，

SELECT DISTINCT Y FROM **K**

からは集合 $\{a, b\}$ を得る．

そこで，K をデータベースを表すファジィ関係と考えよう．実際のファジィデータベースは可能性分布などの複雑なデータを含んでいるが [29]，ここでは，$X_1 \times \cdots \times X_h$ のファジィ集合という単純な場合に限って考える．これに射影演算 $\pi_I(K)$ を考える場合，やはり，$f(x) = \pi_I(x) = (x_{i_1}, ..., x_{i_\ell}, \mu)$ (μ はこのタプルのメンバーシップ) から，$f[K], f(K), \mathcal{P}(f[K])$ などが考えられる．一般に K はファジィマルチ集合と考える必要があるので，

SELECT ALL $i_1, ..., i_\ell$ FROM K

は $f[K]$ に対応するが，

SELECT DISTINCT $i_1, ..., i_\ell$ FROM K

に対応する機能は $f(K)$ ではなく $\mathcal{P}(f[K])$ である．この場合，計算は $\mathcal{P}(f[K])$ あるいは $f(\mathcal{P}(K))$ によるべきである．これと同じ結果を得るとしても，(3.21) を用いた $\mathcal{P}(f(K))$ は計算量が多く効率が劣る．

例 4.4 ファジィ関係スキーマ $\mathbf{K}$(C,Y,M) を考える．ただし，M はメンバーシップに対応するものとする．従って，M は SELECT 文で特に指定する必要はない．タ

プルの集合を

$$
\begin{aligned}
K &= \{(C_1, a, 0.2), (C_2, a, 0.1), (C_1, b, 0.5), \\
&\quad (C_3, a, 0.3), (C_2, b, 0.5), (C_3, b, 0.2)\}.
\end{aligned}
$$

とする．SELECT ALL Y FROM **K** の結果は

$$\{(a, 0.2), (a, 0.1), (b, 0.5), (a, 0.3), (b, 0.5), (b, 0.2)\}$$

となり，SELECT DISTINCT Y FROM **K** の結果は

$$\{(a, 0.3), (b, 0.5)\}$$

となる．

ファジィデータベースにおけるマルチ集合の扱いはこれだけにとどまるものではないが，より詳細な議論は別の機会にゆずりたい．

注意. Yager [30] はクリスプなデータベースの射影からクリスプマルチ集合が生じる例に言及しているが，ファジィデータベースを論じてはいない．

4.4 おわりに

この章では、ラフ集合と関係データベースに対するマルチ集合とファジィマルチ集合の応用について議論した。

関係データベースのように，意識的にマルチ集合を用いている分野もあるが，様々なアルゴリズムにおいて implicit に用いられている場合を含めれば，マルチ集合の適用範囲は大変広いと思われる．

ラフ集合への応用では，$f[A]$ に基づく単純な計算法によって $R^*[A]$ を定義し，上近似を求めている．命題 8 が示しているように，$R^*[A]$ は上近似と下近似の両方を含んでいる．また，ここで述べたラフ集合の扱いは入門的であるので，さらに議論を進める余地が多い．

データベースへの応用では，実際に $f[A]$ が現れる例を示した．データベースに関連した分野として，情報検索への応用 [15] も考えられるが，将来の研究課題としたい．