ラフ集合の概念による推論を用いた設計支援

森　典彦, 高梨　令
デザイン学科

Design Support System based on Reasoning with the Conception of Rough Set Theory

Norihiko MORI, Rei TAKANASHI
Department of Design
(Received October 31, 1996 ; Accepted January 10, 1997)

1.　背景と目的

アンケート調査などで設計の対象について，その属性と，性能やイメージなどの評価との関係のデータを得て(図1)，それを分析して原因となる属性と，結果となる評価との因果関係を知識として獲得しようという試みが,設計の具体化計画に役立たせるためによくなされる。

それはその知識があればその対象の新しい設計を考えるとき，性能やイメージなどについての目標がコンセプトとして与えられているとして，設計の試案がコンセプトに適合しているかどうか，つまり目標を満たすものかどうかという評価を，その試案のもつ属性から推論できるからである。

その分析に使われる手法としては，因果関係を線形で近似できる場合は線形回帰モデルがよく用いられる。それは分析で得られる回帰式を使って設計試案の評価を推論できるからであるが，それ以外に，得られる回帰係数により，ある評価に対して寄与の大きい属性を知ることができ，どのような設計をすればよいかを推論できるからである。こうして設計の具体化にとって有用である。

しかし設計やデザイン問題では因果関係は複雑で線形で近似できない場合も多い。属性が形態要素で結果がバランスなどのイメージの場合がその一例である。それは対象の形を形態要素に分解して，各要素の状態の線形結合すなわち個々の要素の積み上げによって全体の形のバランスを説明することはできず，全形態要素の相互の複雑な関係でバランスがどうなるかがきまると考えられるからである。

そのような因果関係のためのモデル，つまり複雑な非線形関係のための優れたモデルとしてニューラルネットワークがあるが，設計試案がコンセプトに適合するかどうかの評価は推論できるが設計の具体化には役立ちにくい。それは複数の属性がからみ合って結果に寄与するために，線形式における回帰係数のように個々の寄与という概念がなく，したがって設計のどこをどうすればいいかを明示的に知ることはできないからである。非線形ゆえの当然のことである。

しかしその場合でも結果に寄与する属性は常に全属性とは限らないであろうことは想像できる。そうであれば少しでも少ない数の属性によって結果を説明するモデルを作り，結果としての目的のために，個々の属性ではないが，いくつか少数の属性を「同時に」どのようにすればよいかを知ることができ，設計の具体化に大いに役立つはずである。そのためにラフ集合の概念が示唆を与える。

本論の目的は，このような背景のもとで，ラフ集合の概念と，それに基づくブール演算を使うことによって，非線形で複雑な因果関係の中から少しでも単純化されたルールを見出し，設計の具体化計画に有用な知識を得ようとするものである。ただし因果関係はカテゴリカルなデータで表されているものとする。

2.　ラフ集合の基本概念[1)2)3)]

現象の因果関係や意味の包含関係を論理式で記述し形式化する試みは17世紀 G. W. Leibniz に始まり，記号論理学として発達した。そして論理数学としてばかりでなく情報工学としても有用な分野を築いてきた。しかし人間の情報処理は0か1かの2値で表されるはっきりしたものばかりではなく，はっきりしないが多分とか，ある程度はとか，不正確だがおおざっぱにいえばとかの情報も処理している。そのための形式化として確率論，様相論理，ファジイ理論などが発達してきたが，その線上にあるものとしてラフ集合論が生まれた。上記の情報例でいえば，不正確だがおおざっぱにいえば，に相当する。ラフ集合の概念は1982年にポーランドの Zdzislaw Pawlak によって提唱された。

	原因		結果
	対象の属性 $\boldsymbol{x}$		評価 $\boldsymbol{y}$
サンプル	$(x_1)_1, (x_2)_1, \cdots, (x_j)_1$	→	y_1
	$(x_1)_2, (x_2)_2, \cdots, (x_j)_2$		y_1
	・ ・ ・		・
	・ ・ ・		・
	$(x_1)_p, (x_2)_p, \cdots, (x_j)_p$		y_p

図1　設計計画のための推論用データ

ラフ集合の基本概念は分類と近似である。複数のアイテム（属性）に対して複数の対象があり，各対象はアイテムごとに固有のカテゴリーに分類されているものとする。アイテムの一つは結論とし，対象は結論によって識別される。このとき対象の間で同一カテゴリーに分類されるアイテムの集合をRとするとき，Rによって対象を完全に識別できる（結論と符合する）ならばそのことをR定義可能といい，識別不可能ならばR定義不可能という。そしてR定義不可能な対象の集合をRラフ集合という。同類に属することを同値関係にあるというからRは同値関係の集合である。Rの要素数（同値関係の数すなわちカテゴリーが同じであるアイテムの数）を減少すること（また減少したR）を縮約という。定義可能な範囲内でのR要素数最小の縮約は極小縮約とも呼ばれ，極小縮約における結論カテゴリーと他のアイテム・カテゴリーとの関係を極小縮約ルールという。R縮約ルールにおける結論をR下近似といい，極小縮約によって確実に定義される対象の集合である。その縮約は様相論理の概念でいえば必然性の縮約である。またR定義不可能な集合(ラフ集合)をR上近似といい，Rによって対象を完全にはある結論として識別できない（定義できない）がその結論を包含する(識別される可能性のある)集合である。その縮約は様相論理でいえば可能性の縮約である。当然上近似は下近似を包含する。

なお設計関係のデータである図1をラフ集合のデータと見立てた場合，評価が結論で，属性行列も評価ベクトルも要素はカテゴリカルである。

3. 既存のアルゴリズムと本論の考え方

ラフ集合[4]の概念から出発して患者の症状と病気の因果関係についての診断用エキスパートシステムを構築しようという試みがなされてきた[5]。その流れの一つに，因果関係の行列を識別行列に置き換えることでルール発見を容易化する方法[6]や，さらにこれを改良したShan/ZiarkoのアルゴリズムUnknown[7]がある。これらは必然性縮約を求めるほか，演算の途中で好きな確実さでの可能性縮約も求められるところにラフ集合の概念に沿った特徴がある。

	アイテム 1	2	3	4	5	結論 Y
U 1	A	C	E	G	I	1
2	B	C	E	H	I	2
3	A	C	F	H	J	2
4	A	D	F	G	J	2

図2　データ

U	2	3	4
1	A	E	C
	G	G	E
		I	I

↑ or ↓

→ and →

図3　Y＝1の識別行列

U	I
2	B
	H
3	F
	H
	J
4	D
	F
	J

図4　Y＝2の識別行列

ここでShan/Ziarkoのアルゴリズムを紹介する。

まず識別行列を作る。データとして図2が与えられたとする。すなわち対象Uは4個，アイテムは1から5までの5個でそれぞれAとB，CとD，EとF，GとH，IとJの2個ずつのカテゴリーからなり，対象を分類する結論Yは1と2の2個のカテゴリーである。識別行列はY＝1，2のそれぞれにおいて図3，図4のようにそのYに該当する対象を行に，該当しない対象を列に置き，行の対象と列の対象とをデータで比較したとき互いに異なるカテゴリーをもつアイテムの，行の方の対象のカテゴリーを行列の要素として記入したものである。識別行列は結論ごとに作られる。

Shan/Ziarkoのアルゴリズムは，識別行列におけるi行j列の要素をdijとするとき縮約Rを

$$R=\bigwedge_{j}\left\{\bigvee_{i}\ \mathrm{dij},\ \mathrm{dij}\neq\Phi\right\}$$

で求める。ここで論理記号∨はor結合，∧はand結合を表し，Φは空集合を表す。

図3の識別行列，Y＝1についてやってみよう。列のうち空集合を除いて最小の要素数のものから始める。ここではU＝2である。行と行の間の関係はor結合，列と列の間の関係はand結合であることに着目すると，U＝2とU＝3に関する可能性縮約は共通要素Gと，これを除く

U	2･3	4
1	G AE AI	C E I

図5　Y＝1の識別行列からの途中演算

要素の組み合わせでG, AE, AIと求まる。AEはA and Eを示す。この結果とU=4とから図5の行列が得られる。ここでも同様に共通要素をもつ項のandすなわち大なる方のAE, AIと，これを除く項の組み合わせでAE, AI, GC, GE, GIと求まる。これはすべての列についての縮約であるから必然性縮約となっている。

さて，本論の目的に照らして考えると，設計に限らず，一般に何らかの目的があって因果関係のデータから推論のルールを得たいと思ったとき，実験やアンケート調査などで得られるデータは限られる。そのデータを最大限に活用しようとするのでルールは少なくともそのデータの範囲内では確実なもの，つまり必然性あるものとしたい。設計の具体化の場合も同じで可能性は実用上あまり役立たないと思われる。そこで本論でも必然性の縮約のみ求めることにする。

このときブール演算が使えることに注目する。それによりアルゴリズムはずっとすっきりしたものになり，計算機で答えを一気に求めることができる。すなわち各サンプルの属性をand結合，サンプル間をor結合してデータを表し，ブール演算則を用いる。たとえばand結合を×，or結合を＋で表せば

$$A\times A=A^2=A,\ 2A=A,\ A+AB=A(1+B)=A$$

などと単純化がなされ，最終的に極小縮約に至る。たとえば図3のY=1の識別行列に対してブール演算を実行すると

$$\begin{aligned}&(A+G)(E+G+I)(C+E+I)\\&=(AE+GE+AG+G+AI+GI)(C+E+I)\\&=(AE+G+AI)(C+E+I)\\&=AEC+GC+AIC+AE+GE+AIE\\&\quad+AIE+GI+AI\\&=AE+AI+GC+GE+GI\end{aligned}$$

となって同じ結果が得られる。

4.　自動車のイメージへの応用

ラフ集合論における結論と他のアイテムを因果関係と捉え，表1に示す自動車の形態要素を原因としての属性，イメージを結果としての評価とし，表2に示す17個のサンプルについて著者が写真で判定したデータを表3に示す。4個のイメージについての演算で得られたすべての必然性縮約を表4に示す。またある評価カテゴリーを結果とする全サンプルのうちある縮約にあてはまるサンプ

表1　自動車の形態要素とイメージ

1 丸み（線，面，かどの）
　A丸い　B中間　C角ばっている
2 キャビンとボデーの関係
　D分割型　E半融合型　F一体化型
3 ラジエータグリル
　G目立つ　H小さい　Iなし
4 ヘッドランプ
　J丸か曲線で表情あり　K長方形　Lなし
5 ヘッドランプの側面への回り込み
　Mあり　Nなし
6 バンパー
　Pボデーと別色　Qモールつき　R同色
7 ピラーの目立ち度
　Sリヤのみ太い　Tセンターもリヤもやや太い
　Uリヤのみやや太い
8 キャビンの大きさ
　V大きい　W中くらい　X小さい

イメージ　1スポーツ　2パーソナル　3ファミリー　4フォーマル

表2　サンプル

1 Honda Integra	2 Honda Prelude	3 Nissan silvia
4 Eunos 500	5 Toyota Supra	6 Mitsubishi Diamante
7 Nissan Micra	8 BMW 850I	9 Rover 620si
10 Cadillac Eldorado	11 Mercedes Coupe S	
12 Ferrari 456 GT	13 Pontiac Grand AM	
14 Zagato Seta	15 Zagato Bambu	
16 Fiat Tempra	17 Bugatti EB 112	

表3　サンプルの形態要素とイメージ評価

サンプル	形態要素 1	2	3	4	5	6	7	8	イメージ
1	B	D	I	J	N	R	U	X	1
2	A	E	H	J	M	R	S	X	1
3	B	D	H	K	M	R	S	X	1
4	A	E	H	J	M	R	S	W	2
5	A	E	I	J	N	R	U	X	1
6	C	D	G	J	M	Q	U	W	4
7	A	F	H	J	N	P	T	V	3
8	B	D	H	L	M	R	U	W	1
9	B	E	G	K	M	R	S	V	2
10	C	E	G	K	M	Q	S	W	4
11	A	F	G	J	N	R	S	W	2
12	A	E	I	L	N	R	S	X	1
13	B	D	G	J	N	R	S	W	4
14	A	D	I	J	N	R	U	X	1
15	C	E	I	L	N	R	U	X	1
16	C	F	G	K	M	P	U	V	3
17	A	F	H	J	M	R	S	V	2

表4　自動車のイメージに対する形態要素の縮約ルール

縮約とCI (Covering Index)							
Y=1		Y=2		Y=3		Y=4	
BU	2/8	WA	2/4	HN	1/2	CD	1/3
I	5/8	WEH	1/4	P	2/2	CJ	1/3
NU	4/8	WEJ	1/4	T	1/2	CW	2/3
UR	5/8	WER	1/4	VN	1/2	GD	2/3
X	7/8	WHJ	1/4	FC	1/2	GJM	1/3
BDM	2/8	WRJM	1/4	FK	1/2	GUJ	1/3
BH	2/8	WSH	1/4	FMG	1/2	GUW	1/3
DH	2/8	WSJM	1/4	FU	1/2	GWM	2/3
DRM	2/8	WSRM	1/4	FVG	1/2	MDJ	1/3
DSM	1/8	EB	1/4	KU	1/2	MUJ	1/3
KD	1/8	GBK	1/4	VC	1/2	Q	2/3
KH	1/8	GBM	1/4	VU	1/2	WDJ	2/3
NE	3/8	GRE	1/4			WUJ	1/3
UA	2/8	GRK	1/4			CGE	1/3
UE	2/8	GRM	1/4			CKE	1/3
BWM	1/8	KER	1/4			CME	1/3
L	3/8	VB	1/4			CS	1/3
UH	1/8	VE	1/4			GWE	1/3
DA	1/8	VR	2/4			WK	1/3
CR	1/8	VS	2/4			GJB	1/3
NC	1/8	FGJ	1/4			GNB	1/3
		FGN	1/4			GWB	1/3
		FR	2/4			SBJ	1/3
		FS	2/4			SDJ	1/3
		FW	1/4			SNB	1/3
		GA	1/4			SND	1/3
		SNJA	1/4			WBJ	1/3
		FMA	1/4			WBS	1/3
		FMH	1/4			WDS	1/3
		FMJ	1/4			WNB	1/3
		VMA	1/4			WND	1/3
		VMH	1/4				
		VMJ	1/4				

ルの占める割合をその縮約のCovering Index (CI) といい，併せて表4に示す。

表4のうちY=1「スポーティ」に対する縮約を言語表現したものが表5である。スポーティな車をデザインするにはこれらのルールのいずれかを造形に取り入れればよい。

5. 考察と今後の課題

複雑で非線形な関係をブール演算により縮約できることを示した。本論の応用事例のように属性数が8個ぐらいまでならデータを時間かけて眺めていれば縮約は求められるかもしれないが，実際の場では属性数は数十にのぼると思われ，そのときブール演算の威力が発揮される。属性数に対しサンプル数の制約がないことも長所である。

表5　自動車をスポーティイメージにする縮約ルールの言語表現

スポーティイメージにするには		**CI**
I	グリルなしにする	5/8
X	キャビンを小さくする	7/8
L	ヘッドランプなしにする	3/8
リヤピラーのみやや目立たせながら		
UB	全体の丸みを中間にするか	2/8
UN	ヘッドランプの回り込みなしにするか	4/8
UR	バンパーを同色にするか	5/8
UA	全体を丸くするか	2/8
UE	またはキャビンとボデーを半融合にする	2/8
グリルを小さくしておいて		
HD	キャビンとボデーを分割型にするか	2/8
HB	または丸みを中間にする	2/8
キャビンとボデーを半融合にしておいて		
EN	ヘッドランプの回り込みをなしにする	3/8
キャビンとボデーを分割型にし，ヘッドランプを回り込ませておいて		
DMR	バンパーを同色にするか	2/8
DMB	丸みを中間にする	2/8
(CIが2/8未満は省略)		

得られた縮約のうち，属性数の多いものは実際の知識として有用でないので捨てた方がよい。CIの低いものも該当例が少ないという意味で信頼性が低いから避けた方がよい。CIが高くかつ属性数の少ない縮約が最も有用な指針となる。

また必然性のルールといっても得られたデータの中での必然性であって，データにない対象，あるいは新たに設計する上で可能な存在に対しても必然的である保証はない。その意味で必然性縮約も可能性を示すに過ぎず，推測のルールなのである。データのサンプル数が多いほど真の必然性に近づき信頼性が増す。

今後は回帰分析や正準相関分析などとの実務面での適正な使い分けについて整理しておく必要がある。

註

1) W. P. Ziarko (Ed.), Rough Sets, Fuzzy Sets and Knowledge Discovery, Springer-Verlag 1994
2) 中村昭，津本周作，田中博，小林聡，ラフ集合理論とその応用，人工知能学会誌 Vol. 11 No. 2, 1996
3) 中村昭，ラフ集合—その理論と応用，日本ファジィ学会誌 Vol. 8 No. 4, 1996
4) 中村昭他，ラフ集合—その理論と応用，第1回—第2回，数理科学7-12月号，1994.
5) *ibid.*, 第4回.
6) *ibid.*, 第5回.
7) *ibid.*, 第6回.