

STRIM による if-then ルール導出原理とその有効性の検証

Principle on if-then rule induction by STRIM and its confirmation for usefulness

加藤 裕一^{*1}
Yuichi Kato

佐伯 徹郎^{*2}
Tetsuro Saeki

水野 祥太郎^{*1}
Shotaro Mizuno

^{*1} 島根大学 総合理工学部
Interdisciplinary Faculty of Science & Engineering, Shimane Univ.

^{*2} 山口大学 工学部
Faculty of Engineering, Yamaguchi Univ.

This paper proposes an effective reduct method of the decision table from the view of STRIM (statistical rule induction method) and confirms the usefulness by a simulation experiment, comparing with the conventional rough sets method and applying the method for the if-then rule induction problem.

1. はじめに

ラフ集合理論[1]は、決定表と呼ばれるデータベースを if-then ルールで要約する手法として利用される。決定表は、多くのサンプルデータ集合で、各サンプルは通常離散値を取る幾つかの条件属性とそれらの値を条件部、決定属性とその値を結論部とする if-then ルールと見なせる。ラフ集合理論は、このような多くのルールデータを“識別不能”性を利用して簡潔に整理し、要約する。要約表現されたルールは、対象の性質や構造の考察、知識獲得或いは判定問題等に役立つ。このような整理の第一歩は、決定属性に無関係な条件属性を先ず発見して、決定表から除去し同表を簡潔にすることである。この問題は縮約 (reduct) 問題と言われ、従来「無関係性」をどのように定めるか、言い換えれば「決定属性が依存する条件属性集合」を如何に定めるかにより、多くの縮約法が報告されている[2-6]。

本論文は従来のラフ集合理論並びに縮約手法の基本的な考え方を整理して、シミュレーション実験で“識別不能(識別)”性に基づく従来の縮約法の問題点を指摘する。この上で本論文は、「決定属性が依存する条件属性集合」を統計的に判定する新たな縮約法とこれに基づくルール導出法を提案する。具体的には文献[7-9]の STRIM(statistical rule induction method)の考察に従い、サンプルデータは、母集団から得られた各条件属性の確率変数の実現値が、ルールにより決定属性値に変換されたものと見なす。この観点に立てば縮約問題は、サンプルデータを利用した各条件属性の確率変数と決定属性の確率変数との統計的独立検定問題となる。本論文では、両属性に成立する大域的縮約と各 if-then ルール毎に成立する局所的縮約の 2 つの縮約法を提案している。シミュレーション実験で両統計的縮約法の正当性を検証すると共に、ルール導出問題に適用してその有効性を示す。

2. 従来のラフ集合理論と縮約

ラフ集合は、与えられた決定表 S に埋もれた評定構造等を if-then ルール形式で導出する手法として利用される。ここで S は $S=(U, A=C \cup \{D\}, V, \rho)$, $U=\{u(i)|i=1, \dots, N=|U|\}$ はサンプル集合, A は属性集合, $C=\{C(j)|j=1, \dots, |C|\}$ は条件属性集合, $C(j)$ は条件属性, D は決定属性, $V=\bigcup_{a \in A} V_a$ は属性値の集合, $\rho: U \times A \rightarrow V$ は情報関数, と形式表現される。例えば $a=C(j) \in C(j=1, \dots, |C|)$ のとき $V_a=\{1, 2, \dots, M_{C(j)}\}$, $a=D$ とすると

$V_a=\{1, 2, \dots, M_D\}$ となる。Table1 は $|C|=6$, $M_{C(j)}=6$, $M_D=6$, $\rho(u(1), C(1))=5$, $\rho(u(2), C(2))=5$, ... の一例である。

ラフ集合は, U に対して次の識別不能関係: $I_C=\{(u(i), u(j)) \in U^2 \mid \rho(u(i), a)=\rho(u(j), a), \forall a \in C\}$ に着目する。この関係は同値関係であり, U の商集合: $U/I_C=\{[u_i]_C \mid i=1, 2, \dots\}$ を生起させる。ここで $[u_i]_C=\{u(j) \in U \mid (u(j), u_i) \in I_C, u_i \in U\}$ である。 $[u_i]_C$ は u_i を代表元とする同値類である。今 $\forall X \subseteq U$ を取れば, X は同値類を利用して, $C_*(X) \subseteq X \subseteq C^*(X)$ と近似できる。ここで,

$$C_*(X)=\{u_i \in U \mid [u_i]_C \subseteq X\}, \quad (1)$$

$$C^*(X)=\{u_i \in U \mid [u_i]_C \cap X \neq \emptyset\} \quad (2)$$

で, $C_*(X)$, $C^*(X)$ は夫々 X の C による下近似, 上近似と呼ばれる。通常 $(C_*(X), C^*(X))$ は X の C によるラフ集合と呼ばれる。更に $X=D_d=\{u(i) \mid (\rho(u(i), D)=d)\}$ とすれば, 定義より $C_*(X)$ は確実に $D=d$ となる集合であり, $C^*(X)$ は $D=d$ となる可能性のある U の部分集合である。従って $C_*(X)$ 或いは $C^*(X)$ から夫々, 必然性或いは可能性 if-then ルールが導出される。

Table 1 Example of decision table

U	(C (1) C (2) C (3) C (4) C (5) C (6))	D
1	563242	3
2	256124	6
3	116226	1
4	416646	6
...	...	...
N-1	151252	2
N	513135	4

従来のラフ集合は、決定属性が依存しない或いは無関係な条件属性を検出して決定表から除去し、決定表を整理することとを議論する。整理された決定表はルール導出負荷を軽減すると共に、ルールの特徴を明瞭化する。この問題は、縮約 (reduct) 問題と呼ばれ、従来様々な縮約手法が報告されている。例えば、 $B \subseteq C$ に対して、次の 2 つの条件:

(i) $B(D_d)=C_*(D_d)$, $d=1, 2, \dots, M_D$,

(ii) $(B-\{a\})_*(D_d)=C_*(D_d)$, $d=1, 2, \dots, M_D$, となる $a \in B$ が存在しないとき, B は下近似保存縮約と呼ばれる。同様に, 上近似保存縮約なども考えられる。

代表的な縮約手法として LEM1 アルゴリズム[2]や識別行列法[3]がある。LEM1 アルゴリズムは, (i)(ii)を原則に従い実施する例である。 $|C|$ の増加に伴い, 計算量が指数関数的に増加す

連絡先: ^{*1}ykato@cis.shimane-u.ac.jp, ^{*2}tsaeki@yamaguchi-u.ac.jp

る。識別行列法は、次の(i, j)成分 δ_{ij} を持つ $N \times N$ の対称行列を作成する。

Table 2 Hypothesis with regard to the decision attribute value

仮説 1	$u^0(i)$ は if-then ルールの条件部に当てはまり、結論部の決定属性値 $u^0(i)$ が一意に定まる (一意決定可能データ)
仮説 2	$u^0(i)$ は if-then ルールのどの条件部にも当てはまらず、 $u^0(i)$ をランダムに決めざるを得ない (無関係データ)
仮説 3	$u^0(i)$ は複数の if-then ルールの条件部に当てはまり、一意に $u^0(i)$ を決定出来ず、複数の該当ルールの一つをランダムに選んで $u^0(i)$ を決定せざるを得ない (矛盾データ)

$$\delta_{ij} = \{a \in C \mid \rho(u(i), a) \neq \rho(u(j), a)\}; \exists d \in D, \rho(u(i), d) \neq \rho(u(j), d) \text{ かつ } \{u(i), u(j)\} \cap \text{Pos}(D) \neq \emptyset \text{ のとき,} \\ = * ; \text{その他 } (U - \text{Pos}(D))$$

ここで $\text{Pos}(D) = \bigcup_{d=1}^M C_*(D_d)$, * は don't care の意である。この上で次の演算を実施する。

Table3 Results of test for independence by Bootstrap method

X ²		C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
AVE		341.69	360.59	313.27	327.40	24.79	25.23
S.D.		37.35	35.91	35.31	34.55	7.04	6.06
Max		451.19	471.25	412.19	418.39	47.83	40.61
Min		270.7	285.91	232.56	255.42	10.93	8.21
p-value	Max	9.44E-80	6.84E-84	9.87E-72	5.28E-73	3.91E-03	2.52E-02
	Min	4.29E-43	3.98E-46	1.45E-35	4.59E-40	9.93E-01	9.99E-01

$$F^{\text{reduct}} = \bigwedge_{i,j:i < j} \vee \delta_{ij}, \quad (3)$$

$\bigwedge$ は連語、 $\bigvee$ は選語である。 F^{reduct} は識別関数と呼ばれる。 δ_{ij} は $u(i)$, $u(j)$ の少なくとも一方が $\text{Pos}(D)$ の要素で、決定属性値が異なる原因の条件属性 (識別できる条件属性) を要素としており、これが全ての要素で成立する条件を求めているのが (3) 式である。逆に識別不能な要素には無関係である。この演算結果の最簡形が reduct となる。識別行列法も LEM1 と同様な問題を持つ。詳細なアルゴリズム等は文献[2][6][10][11] 参照。

3. 従来の縮約手法の追試実験

ここでは従来の基本的な縮約法である LEM1[2] 及び識別行列法[3] の性能を、シミュレーションデータを利用して具体的に追試した結果を示す。文献[7-9] では、ランダムに発生させた条件属性値の組を予め設定したルール集合と Table2 に示す仮説を利用して、決定属性値に変換するデータ発生モデルを提案している。このモデルを利用して発生したデータに両縮約法を適用した。具体的には決定表の条件属性数とその属性値数及び決定属性値数は Table1 の設定とした。設定ルールは、

$$R(d): \text{if } Rd \text{ then } D=d \ (d=1, \dots, M_D=6), \quad (4)$$

$Rd = (C(1)=d) \wedge (C(2)=d) \vee (C(3)=d) \wedge (C(4)=d)$ とした。この上で一様乱数を利用してサンプル $u(i)$ の条件属性値 $u^C(i) = (v_{C(1)}(i), v_{C(2)}(i), \dots, v_{C(C)}(i))$ を発生し、上記設定ルールと Table2 の仮説を利用して、標本 $u(i)$ の決定属性値 $u^D(i) (i=1, 2, \dots, N)$ を決定し

た。 $N=10000$ として、LEM1、識別行列法を Table1 に適用して縮約実験した。予め設定したルールから $F^{\text{reduct}} = C(1) \wedge C(2)$

Table4 Example of if-then rules for local reducts

Rule No.	specified rule: $(C(1)C(2), \dots, C(6)D)$
1	(1100001)
2	(0011001)
3	(0000222)
4	(2200002)
5	(0033003)
6	(0000333)
7	(4400004)
8	(0044004)
9	(0000555)
10	(5500005)
11	(0066006)
12	(0000666)

$\wedge C(3) \wedge C(4)$ となり、決定属性に無関係な条件属性は $C(5)$, $C(6)$ となるはずである。しかし両手法は $F_{LEM1}^{\text{reduct}} = F_{Disc}^{\text{reduct}} = C(1) \wedge C(2) \wedge C(3) \wedge C(4) \wedge C(5) \wedge C(6)$ を導出した。ここで F_{LEM1}^{reduct} , F_{Disc}^{reduct} はそれぞれ LEM1 と識別行列法の意である。更に、Table1 を新たに発生して、併せて 3 回の実験した結果、LEM1、識別行列法とも同一結果となった。

この原因は、従来のラフ集合が「識別不能性」或いは「識別性」を基盤としていることである。従来の識別法は、サンプルデータに潜む必然的な差異 (一意決定性 (Table2 参照)) と偶然性 (矛盾・無関係データ) によるものとを整理して識別出来ない。これが従来の reduct 法の問題点である。

Table5 Results of test for independence by Bootstrap method using sample data set generated by rules in Table4

χ^2		C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
AVE		209.56	206.35	244.57	234.94	252.45	219.59
S.D.		29.38	29.08	30.14	30.72	27.41	28.83
Max		293.56	290.93	318.17	317.52	310.48	314.13
Min		140.01	147.43	185.15	162.5	186.58	153.62
p-value	Max	1.17E-47	3.95E-47	1.33E-52	1.80E-52	4.72E-51	8.69E-52
	Min	5.71E-18	2.51E-19	2.14E-26	4.03E-22	1.14E-26	1.81E-20

4. 条件属性の統計的縮約法の提案

4.1 大域的縮約法

3. で述べたように、従来のラフ集合の縮約手法は、基本的に矛盾・無関係データへの対応性を持たない。そこで、文献[7-9] で考察した STRIM の観点から、この縮約問題を考察する。STRIM は決定表を母集団からの一標本データ集合と見なしている。確率・統計モデルに従えば、 $u(i) = (v_{C(1)}(i), v_{C(2)}(i), \dots, v_{C(C)}(i), u^D(i))$ は母集団を観測する確率変数 $A = (C(1), C(2), \dots, C(C), D) = (C, D)$ の実現値である (誤解の生じない範囲で属性名を確率変数としても以降利用する)。

今 C に関して次の確率モデルを設定する。任意の j に対して、 $P(C(j) = v_{C(j)}(k)) = p(j, k)$ $\sum_{k=1}^{M_{C(j)}} p(j, k) = 1$, 任意の $j_1 \neq j_2$ に対して $P(C(j_1) = v_{C(j_1)}(k_1), C(j_2) = v_{C(j_2)}(k_2)) = p(j_1, k_1) p(j_2, k_2)$ 即ち、

C(j1)と C(j2)は独立とする。(4)の設定ルールで言えば、 $C=(1,1,2,3,4,5)$ (以降(112345)と簡便に記す)であれば、Table2の仮説1により $P(D=1)=1$ である。 $C=(123456)$ であれば、仮説2により $P(D=1)=1/M_D=1/6$ である。 $C=(112256)$ であれば、仮説3により $P(D=1)=1/2$ である。一般にDは、Cの実現値と母集団の if-then ルール並びに仮説から定まる確率変数で、

$$P(D, C) = P(D|C)P(C) = (v_{c(1)}(i), v_{c(2)}(i), \dots, v_{c(|C|)}(i)) \quad (5)$$

Table 6 Examples of contingency table for local reducts

(N=3000, df=5): D=1 vs. C

D=1 U(D=1) =486	C					
	1	2	3	4	5	6
1	141	122	137	141	78	88
2	64	67	89	80	95	74
3	82	86	59	82	74	79
4	58	77	53	52	93	101
5	79	58	83	75	71	64
6	62	76	65	56	75	80
X^2	52.83	35.01	66.42	53.16	2.99	13.33
p-values	3.64E-10	1.5E-06	5.68E-13	5.00E-10	0.702	2.05E-2

となる。ここで、 $P(D|C)$ はCに依存してDを取る確率で、if-thenルールに依存する。特別な場合、 $C(j)(j=1, \dots, |C|)$ がルール条件部に現れなければ、Dの実現値はC(j)に依存しない。従って $P(D, C(j)) = P(D|C(j))P(C(j)) = P(D)P(C(j))$ が両者の「無関係」性である。そこで確率・統計的立場から、標準データを用いて両者の「無関係」性即ち、独立性は次の2つの仮説：帰無仮説 $H0(j): C(j)$ とDは互いに独立である、対立仮説 $H1(j): C(j)$ とDは互いに独立ではない、の下で検定すればよい。仮説検定の性質(第2種の過誤は統制出来ない)から、ここでは $H0(j)$ が棄却されたC(j)のみで決定表を構成することを提案する。即ち、 $H0(j)$ が棄却されなかったC(j)を決定表から削除することで、決定表を縮約することを提案する。

具体的には、カテゴリー値を取る2つの確率変数の独立性を検定する標準的な手法として、統計分野でよく知られた χ^2 分布検定を利用する。C(j)の実現値 $v_{C(j)}(i) (\in \{1, 2, \dots, M_{C(j)}\})$ とDの実現値 $u_D(i) (\in \{1, 2, \dots, M_D\})$ の $M_{C(j)} \times M_D$ の分割表は、適切な検定条件を満たせば、 χ^2 値は自由度 $df=(M_{C(j)}-1) \times (M_D-1)$ の χ^2 分布する[12]。Table1で $N=10000$ データから $N_B=3000$ 個無作為に抽出して、新たな決定表を作成する。この

Table 7 Arrangement of Car Evaluation data set of UCI

unified attribute value	C(1):buying	C(2):maint	C(3):doors	C(4):person	C(5):lug boot	C(6):safety	D:class (freq.)
1	vhigh	vhigh	2	2	small	low	unacc (1210)
2	high	high	3	4	med	med	acc (383)
3	med	med	4	more	big	high	good (89)
4	low	low	5more	-	-	-	vgood (65)

決定表に本縮約法を適用する実験を併せて $N_r=100$ 回実行した(Bootstrap法)。この結果をTable3に整理している。同表には全てのC(i)とD($i=1, 2, \dots, 6$)との χ^2 値の平均・標準偏差、最大

Table 8 Results of global reduct for Car Evaluation data set

	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
X^2	189.24	142.84	10.38	371.34	53.28	479.32
p-value	5.92E-36	2.54E-26	3.20E-01	4.04E-77	1.03E-08	2.39E-100

値(Max)・最小値(Min)及びこの χ^2 値に対するp-値を整理している。同表から、(1)予めDと独立な設定のC(5)並びにC(6)と、独立でないC(1)~C(4)との χ^2 値の変動域は重なっていないこと、(2)両者のp-値には明らかな差異があることが分かる。Dに対して $H0(j)$ が棄却されなかったC(j)を検出して決定表から除去することを、ここでは大域的縮約法と呼ぶ。本提案縮約法の有効性が窺える。

4.2 局所的縮約法

本節では大域的縮約が存在しない場合でも、全ての if-then ルール毎に適用できる局所的縮約法を提案する。(5)を決定属性値 $D=\ell$ について記せば、

$$P(D=\ell, C) = P(D=\ell|C)P(C) = (v_{c(1)}(i), v_{c(2)}(i), \dots, v_{c(|C|)}(i)) \quad (6)$$

となる。従って、4.1と同様にDとC(j)に関する独立性問題を $D=\ell$ とC(j)の独立性問題として、DとC(j)に関する仮説を $D=\ell$ とC(j)に関する仮説として議論すれば、 $D=\ell$ 毎に決定表を縮約することができる。この縮約は $D=\ell$ 毎の if-then ルール導出に利用できる。Table4に新たな設定ルールを示す。この設定ルールは大域的な縮約を持たない例である。このルールに従って、3章と同様にシミュレーションデータを発生し、4.1節の縮約法を確認した結果をTable5に示す。Table4に設定したようにTable5では大域的な縮約属性を検出することは困難である。そこで、一例として $D=1$ となるルール(Table4のRule No.1,2)での縮約確認の考察を行ったものをTable6に示す。設定通り、C(1),...,C(4)は $D=1$ を構成するルールに関与している(独立ではない)ことが明確である。一方C(5),C(6)にはC(1),...,C(4)のp-値と比較して、大きな段差があり、 $D=1$ のルールに明確に関与していると主張できないことが、分かる。従って、 $D=1$ となるルールは、C(1)~C(4)の条件属性の論理式によって構成できる知見が得られる。更に、C(1)=1, C(2)=1, C(3)=1, C(4)=1の度数が多いことから、これらの条件属性値を用いてルールを構成することも示唆している。この知見はTable4で設定したルールと一致している。他の $D=2, \dots, 6$ についても同様である。

Table 9 Results of local reduct for Car Evaluation data set

		C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
D=1	X^2	22.94	19.24	2.59	111.00	3.31	113.31
	p-value	4.16E-05	2.44E-04	4.62E-01	7.88E-25	1.22E-02	1.59E-26
D=2	X^2	11.77	10.77	3.19	192.06	0.52	194.25
	p-value	8.21E-03	1.38E-02	3.64E-01	1.53E-42	3.85E-02	6.58E-43
D=3	X^2	84.33	84.33	0.39	34.70	0.26	36.26
	p-value	3.61E-18	3.61E-18	5.42E-01	2.92E-03	8.78E-01	1.34E-03
D=4	X^2	70.20	28.60	4.23	33.08	37.68	130.00
	p-value	3.87E-15	2.72E-06	2.36E-01	6.57E-06	6.53E-09	5.90E-29

5. STRIMによる縮約結果からのルール導出

4.の縮約考察の結果を利用して、文献[13]の"Car Data"に大域的・局所的縮約法を適用した。Table7はこのデータの属性：C(j)($j=1, \dots, 6$)と{D}とその値： $M_{C(j)}=4(j=1, 2, 3)$, $M_{C(j)}=3(j=4, 5, 6)$, $M_D=4$,を整理している。また、各決定属性値毎の度数を示している。このデータ $N=1728$ に大域的並びに局所的縮約法を適用した結果を夫々Table8, 9に示す。Table8はC(3):doorsが縮約できることが分かる。即ち、決定属性classは条件属性doorsに「統計的に独立である」ことを否定できないことが分かる。一方Table9は、 $D=1$ となるルールではC(3),C(5), $D=2$ はC(1),C(2),C(3),C(5), $D=3$ はC(3),C(5), $D=4$ はC(3)を縮約できることが分かる。この結果からもTable8の結果を確認できる。

特に、代表して D=1,4 の分割表(Table6 に対応)を示したのが、Table10 である。同表で4. 2での考察に基づき、D 値に影響する(縮約出来ない)C(j)値の内、最頻度の属性値を斜体・太字で

Table 10 Examples of contingency table and χ^2 test by local reducts

(a) D=1						
$V_{(0)}$	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
1	360	360	326	576	450	576
2	324	314	300	312	392	357
3	268	268	292	322	368	277
4	258	268	292	-	-	-
χ^2	22.94	19.24	2.58	111.00	8.81	118.81
p-value	4.16E-05	2.44E-04	4.62E-01	7.88E-25	1.22E-02	1.59E-26
(b) D=4						
$V_{(0)}$	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)
1	0	0	10	0	0	0
2	0	13	15	30	25	0
3	26	26	20	35	40	65
4	39	26	20	-	-	-
χ^2	70.20	28.60	4.23	33.08	37.69	130.00
p-value	3.87E-15	2.72E-06	2.38E-01	6.57E-08	6.53E-09	5.90E-29

示している。Table6 の知見から、D=1 のルール候補は、C(1)=1, C(2)=1, C(4)=1, C(5)=1, D=4 は C(1)=4, C(2)=4, C(5)=3, C(6)=3 の組み合わせから生成できることが分かる。この知見から、幾つかのルール候補を作成したのが Table11 である。同表から D=1 となるルールは、Rule No.=1・2 が妥当と判断される。D=4 となるルールは、Rule No.=13 辺りが妥当と思われる。詳細な考察は当日報告する。

6. おわりに

本論文では、決定表を縮約して、簡潔化した決定表から if-then ルールを導出する手法を考察した。具体的にはシミュレーションデータを利用して、従来の縮約法は殆ど対応できないことを示した。この上で著者らが従来提案している STRIM の考え方に従って、新たな統計的縮約法を提案した。特に局所的縮約法は決定属性値毎に縮約できる上に、決定値属性値が依存する条件属性値を推定できることを示し、if-then ルール推定に利用できることを示した。この有効性をシミュレーションデータ並びに UCI データに適用して検証した。

参考文献

- [1] Z.Pawlak: Rough sets, Internat. J. Inform. Comput. Sci.,

Vol.11, No.5, pp.341-356 (1982).

- [2] J.W.Grzymala-Busse: LERS- A system for learning from examples based on rough sets. In Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory, ed. By R. Słowiński, Kluwer Academic Publishers, 3-18(1992).
- [3] A.Skowron and C.M.Rausser: The Discernibility Matrix and Functions in Information Systems, in: R. Słowiński (ed), Intelgent Decision Support, Handbook of Application and Advances of Rough Set Theory, Kluwer Academic Publishers, 331-362(1992).
- [4] Z.Pawlak: Rough set fundamentals; KFIS Autumn Coference Tutorial, pp.1-32(1996).
- [5] D.Ślęzak: Various approaches to reasoning with frequency based decision reducts: A survey, in L.Polkowski, S.Tsumoto and T.Y.Lin(eds):Rough Set Method and Applications, Physical-Verlag, pp.235-285(2000).
- [6] Y.G. Bao, X.Y. Du, M.G. Deng and N. Ishii: An Efficient Method for Computing All Reducts, Transsaction of the Japanese Society for Artificial Intelligence, Vol.19, No.3, pp.166-173(2004).
- [7] 水野祥太郎, 加藤裕一, 佐伯徹郎: 統計的検定法を用いた決定表からのルール導出法; システム制御情報学会, Vol.26, No.8, pp.297-305 (2013).
- [8] Y. Kato, T. Saeki and S. Mizuno: Studies on the Necessary Data Size for Rule Induction by STRIM, In P. Lingras et al. (Eds.): RSKT 2013, LNAI Vol.8171, pp.213-220 (2013).
- [9] 加藤裕一, 佐伯徹郎, 水野祥太郎: STRIM によるルール導出原理と適用範囲に関する考察システム制御情報学会, Vol.27, No.10, pp.385-394 (2014).
- [10] 乾口雅弘: ラフ集合による情報の解析; システム制御情報学会, Vol.49, No.5, pp.165-172 (2005).
- [11] 森典彦, 田中英夫, 井上勝雄編: ラフ集合と感性-データからの知識獲得と推論-, KABUNNDO, pp.163-1184, (2004).
- [12] R. E. Walpole, R. H. Myers, S. L. Myers, K. Ye, Probability and Statistics for Engineers and Scientists, Eighth edition, Pearson Prentice Hall, pp.374-377 (2007).
- [13] A. Asuncion and D. J. Newman: UCI Machine Learning Repository, University of California, School of Information and Computer Science, Irvine (2007), <http://www.ics.edu/~mllearn/MLRepository.html>

Table11 Examples of estimated rules by use of the reduct results for Table10

Rule No.	examples of estimated rules (C(1)C(2)C(3)C(4)C(5)C(6)D)	distribution of decision values (n_1, n_2, n_3, n_4)	p-value(χ^2)	accuracy	coverage
1	(0000011)	(576, 0, 0, 0)	3.58E-53(246.59)	1.00	0.476
2	(0001001)	(576, 0, 0, 0)	3.58E-53(246.59)	1.00	0.476
3	(0001011)	(192, 0, 0, 0)	1.03E-17(82.20)	1.00	0.157
4	(1000001)	(360, 72, 0, 0)	6.47E-11(50.43)	0.83	0.296
5	(0100001)	(360, 72, 0, 0)	6.47E-11(50.43)	0.83	0.296
6	(1000011)	(144, 0, 0, 0)	2.61E-13(61.64)	1.00	0.119
7	(0100011)	(144, 0, 0, 0)	2.61E-13(61.64)	1.00	0.119
8	(1001001)	(144, 0, 0, 0)	2.61E-13(61.64)	1.00	0.119
9	(0101001)	(144, 0, 0, 0)	2.61E-13(61.64)	1.00	0.119
10	(0000034)	(277, 204, 30, 65)	2.24E-37(173.49)	0.113	1.000
11	(0000304)	(368, 144, 24, 40)	1.24E-4(20.65)	0.069	0.615
12	(0000334)	(88, 64, 0, 40)	1.85E-39(183.14)	0.208	0.615
13	(4000034)	(52, 33, 20, 39)	1.24E-57(267.21)	0.27	0.600
14	(0400034)	(52, 46, 20, 26)	7.32E-31(143.30)	0.18	0.400
15	(4000334)	(16, 8, 0, 24)	-	0.50	0.369
16	(4403334)	(0, 0, 0, 4)	-	1.00	0.062