



論理簡単化技術のラフ集合への導入について

メタデータ	<p>言語: jpn</p> <p>出版者: 日本知能情報ファジィ学会</p> <p>公開日: 2013-08-22</p> <p>キーワード (Ja):</p> <p>キーワード (En):</p> <p>作成者: 工藤, 康生, 高木, 昇</p> <p>メールアドレス:</p> <p>所属:</p>
URL	http://hdl.handle.net/10258/2202

論理簡単化技術のラフ集合への導入について

その他（別言語等） のタイトル	On an Approximate Calculation Method of a Relative Reduct Based on Classification by Condition Attributes
著者	工藤 康生, 高木 昇
雑誌名	ファジィシステムシンポジウム講演論文集
巻	25
発行年	2009-07
URL	http://hdl.handle.net/10258/2202

doi: info:doi/10.14864/fss.25.0.107.0

論理簡単化技術のラフ集合への導入について

A Note on Introduction of Logic Minimization Techniques into Rough Sets

工藤 康生

Yasuo Kudo

室蘭工業大学

Muroran Institute of Technology

高木 昇

Noboru Takagi

富山県立大学

Toyama Prefectural University

Abstract: We consider introducing logic minimization techniques into rough set theory. Computational complexity of calculating all relative reducts is NP-hard and therefore some heuristic methods are needed for real-time calculation of decision rules from huge size data. Logic minimization is a useful heuristic method for generating minimal decision rules with respect to minimizing the number of rules.

1 はじめに

ラフ集合 [3] による縮約計算および決定ルール抽出は、データマイニングの一手法として注目されている。しかし、決定表に対するすべての相対縮約の抽出等の計算量は NP 困難であるため [6]、大規模データに対して実時間で決定ルール抽出を行うためにはヒューリスティックな手法が必須である。本稿では、論理回路設計における論理簡単化技術 (詳細は [4] 等) を、ヒューリスティックなルール抽出手法としてラフ集合に導入することを検討する。

2 ラフ集合による決定ルール抽出

本節ではラフ集合による決定ルールの抽出について概説する。なお、本節の内容は文献 [2] に基づく。

ラフ集合による決定ルール抽出では、対象とするデータは以下の 4 項組で定義される決定表として表される:

$$(U, C \cup D, V, \rho),$$

ここで、 U は対象の空でない有限集合、 C は条件属性の空でない有限集合、 D は決定属性の空でない有限集合であり、 $C \cap D = \emptyset$ とする。 V は各属性 $a \in (C \cup D)$ の値の集合、 $\rho: U \times (C \cup D) \rightarrow V$ は対象 x の属性 a での値 $\rho(x, a) \in V$ を表す関数である。

決定表の例を表 1 に示す。表 1 は議論の対象となる要素の集合 $U = \{x_1, \dots, x_6\}$ 、条件属性集合 $C = \{c_1, \dots, c_4\}$ 、決定属性集合 $D = \{d\}$ などで構成される。 $\rho(x_i, d) = \text{yes}$ となる要素の集合を決定クラス D_{yes} とすると、決定クラス $D_{\text{yes}} = \{x_4, x_5, x_6\}$ が得られる。同様に、 $D_{\text{no}} = \{x_1, x_2, x_3\}$ も得られる。

各決定クラス D_i に対して、識別不能関係 R_A による下近似 $\underline{A}(D_i)$ を以下の通り定義する:

表 1: 決定表の例

U	c1	c2	c3	c4	d
x1	1	0	0	0	no
x2	0	0	0	1	no
x3	0	1	1	0	no
x4	1	1	0	0	yes
x5	1	0	0	1	yes
x6	0	1	0	1	yes

$$\underline{A}(D_i) = \{x \in U \mid [x]_A \subseteq D_i\},$$

ここで、 $[x]_A$ は R_A による同値類である。

決定表の相対縮約とは、識別可能なすべての対象をその決定クラスに正しく分類するために必要となる最小限の条件属性の集合である。形式的には、決定表の相対縮約を、すべての決定クラス D_i ($i = 1, \dots, m$) に対して以下の 2 条件を満たす条件属性の部分集合 $A \subseteq C$ として定義する:

1. $\underline{A}(D_i) = \underline{C}(D_i)$, 2. $\underline{B}(D_i) \neq \underline{C}(D_i)$, $\forall B \subset A$,

ここで、 $B \subset A$ は集合 B が集合 A の真部分集合であることを意味する。例として、表 1 の相対縮約は $\{c_1, c_2, c_4\}$ である。一般的に、決定表に対する相対縮約は複数個存在しうる。

相対縮約を用いることで、条件属性の値に対する決定属性の値を表す決定ルールを、より簡略化して表すことができる。例えば、相対縮約 $\{c_1, c_2, c_4\}$ を用いると、決定クラス $D_{\text{yes}} = \{x_4, x_5, x_6\}$ に関する以下の 3 個の決定ルールが得られる。

$$(1) (c_1 = 1) \wedge (c_2 = 0) \wedge (c_4 = 1) \rightarrow (d = \text{yes}),$$

$$(2) (c_1 = 1) \wedge (c_2 = 1) \wedge (c_4 = 0) \rightarrow (d = \text{yes}),$$

$$(3) (c_1 = 0) \wedge (c_2 = 1) \wedge (c_4 = 1) \rightarrow (d = \text{yes}).$$

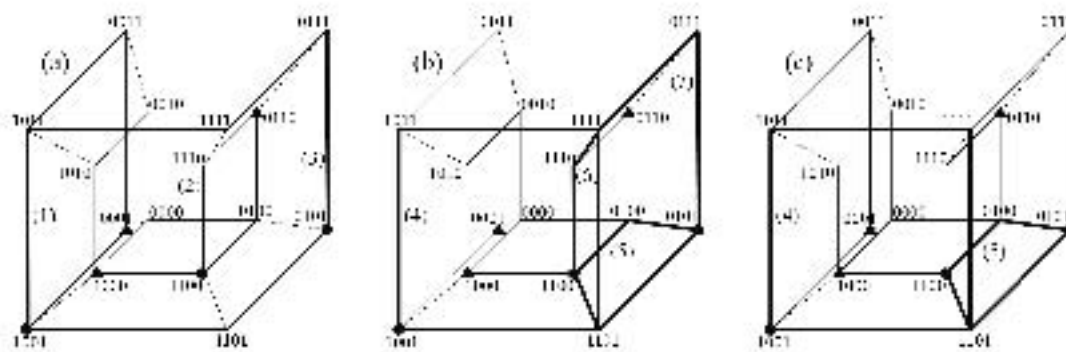


図 1: 決定クラス D_{yes} に関する決定ルールのキューブによる表現: (a) 相対縮約による決定ルール, (b) 決定行列による極小決定ルール, (c) 論理簡単化による決定ルール

条件部の長さが極小となるすべての決定ルール (以下, 極小決定ルールと呼ぶ) を抽出する方法として, 決定行列 [5] を用いる方法が知られている. 例として, D_{yes} からは以下の 4 個の極小決定ルールが得られる:

- (4) $(c1 = 1) \wedge (c4 = 1) \rightarrow (d = yes)$,
- (5) $(c2 = 1) \wedge (c3 = 0) \rightarrow (d = yes)$,
- (6) $(c1 = 1) \wedge (c2 = 1) \rightarrow (d = yes)$,
- (7) $(c2 = 1) \wedge (c4 = 1) \rightarrow (d = yes)$.

3 ラフ集合への論理簡単化技術の導入

論理回路設計における最小化技術は, ラフ集合によるルール抽出と密接に関連している. 論理回路設計では, 設計する回路の大きさを極力小さくするために, 回路を記述する論理関数の積項 (リテラルの積) の個数を減らすことが重視される. また, ヒューリスティックな手法を用いることで, 大規模なデータに対しても実時間で近似解を求めることが可能である [1]. これをラフ集合の枠組みで考えると, ラフ集合に論理簡単化技術を導入することは, 以下の制約条件を満たす近似解を求めることに対応すると考えられる:

- ・与えられた決定クラスに含まれるすべてのデータを, 可能な限り少数の極小決定ルールで被覆する
- ・他の決定クラスに含まれるどのデータも被覆しない

表 1 の決定表および D_{yes} に関する決定ルールを, 4 次元超立方体上に表示した例を図 1 に示す. 4 次元超立方体の各頂点は条件属性の値の組合せを表す積項に対応し, 例えば左下の “1001” は $(c1 = 1) \wedge (c2 = 0) \wedge (c3 = 0) \wedge (c4 = 1)$ を表す. また, は D_{yes} のデータ, は D_{no} のデータであることを意味する. 図 1(a) において, 相対縮約から得られた決定ルール (1)~(3) は, 太線で示された 4 次元超立方体の辺に対応する. 一方, 図 1(b) に示すように, D_{yes} の 4 個の極小決定ルールは, 太線

で囲まれた面に対応する. これに対し, 図 1(c) に示すように, 論理簡単化による決定ルールでは, (4) と (5) の 2 個の決定ルールで, D_{yes} のすべてのデータを被覆できる. これらの例から, 決定ルールに用いる条件属性の種類を属性縮約によって削減することは, 必ずしもルール数の削減をもたらさないことが示唆される.

4 おわりに

本稿では, 論理簡単化技術をラフ集合に導入することで, 極小決定ルールの個数を極力減らすヒューリスティックなルール抽出を検討した. 論理簡単化によるルール抽出の実装および検証等が今後の課題である.

参考文献

- [1] R. K. Brayton et al.: *Logic Minimization Algorithms for VLSI Synthesis*, Kluwer, 1984.
- [2] 乾口 雅弘: ラフ集合と決定表の解析, 森, 田中, 井上編, ラフ集合と感性~データからの知識獲得と推論~, 163~181, 海文堂出版, 2004.
- [3] Z. Pawlak: Rough Sets, *Int. Journal of Computer and Information Science*, **11**, 341~356, 1982.
- [4] 笹尾 勤: 論理設計 —スイッチング回路理論—, 近代科学会, 1995.
- [5] N. Shan and W. Ziarko: Data-Based Acquisition and Incremental Modification of Classification Rules, *Computational Intelligence*, **11**(2), 357~370, 1995.
- [6] A. Skowron and C. M. Rauszer: The discernibility matrix and functions in information systems, *Intelligent Decision Support: Handbook of Application and Advance of the Rough Set Theory*, R. Słowiński (ed.), Kluwer, 331~362, 1992.

連絡先

工藤 康生
〒 050-8585 北海道室蘭市水元町 27-1
室蘭工業大学しくみ情報系領域
Tel: 0143-46-5469, Fax: 0143-46-5499
E-mail: kudo@csse.muroran-it.ac.jp