

斜面崩壊危険度診断事例からのラフ集合論と遺伝的アルゴリズムによる知識獲得手法

関西大学大学院	総合情報学研究科	学生会員	原川 浩一
関西大学	総合情報学部	正 会 員	広兼 道幸
関西大学	総合情報学部	正 会 員	田中 成典
関西大学	総合情報学部		木村 寿夫

1. はじめに

知識獲得に関する研究の目的は、エキスパートシステムに代表される、知識ベースシステムを構築する際に必要となる知識の獲得手法の開発である。現在の知識獲得研究における最大の関心事は、データベースなどに蓄えられた膨大なデータから、如何に知識を抽出するかという事である。ラフ集合論は、近年このような知識獲得の分野で注目されている手法のひとつである。ラフ集合論の概念を用いることによって、例えば、土木技術の専門家による斜面の崩壊危険度診断事例から、矛盾を含む決定規則や不要な判定要因を取り除き、簡潔で整合性のあるルール型知識を取り出すことができることが知られている¹⁾。このようにして抽出した知識を利用することで、より少ない項目だけでの診断が可能になるだけでなく、最後まで残った要素を検証することで、専門家が実際にどの項目を重視しているのかなど、生のデータからでは容易に知ることのできない新しい知識を得られる可能性も考えられる。しかし、ラフ集合の概念を用いたこれらの知識獲得手法では、解の探索が組合せ問題となるため、扱うデータが大きくなればなるほど計算時間もまた膨大なものになることが問題として挙げられている。

そこで本研究では、専門家による斜面崩壊危険度診断事例を用い、そこから専門家の経験的知識を獲得するための手法として、ラフ集合論を基本として用いた。さらに、計算時間の問題を解決するための手法として組合せ問題における解の大域的探索を得意とする遺伝的アルゴリズム (GA: Genetic Algorithm) に着目し、これらの手法を組み合わせた知識獲得手法を提案した。

2. ラフ集合論と GA の概要

人間の知的活動は、その対象をいくつかの属性に従って類別することによって成り立っているといえる。対象物がこの類別に対して同じであれば、それらの対象物は識別不能であるとして取扱い、推論・決定・学習などを行っている。ラフ集合論は、このような人間の情報処理活動における類別能力に着目して提唱された概念である²⁾。本研究では、専門家による斜面崩壊危険度診断事例を、その判定要因 (条件属性) と危険度 (決定属性) からなる決定表として扱い、そこからラフ集合の基礎概念のひとつである識別不能性・類別の考え方に基づいて不要な判定要因や判定要因に対する項目を削除し、決定表を縮約化することで診断に必要な最小限のルール型知識 (以下、極小決定アルゴリズム) を獲得した。一方、GA は生物の遺伝と進化の様子を、Darwin の自然選択説 (Natural Selection) に基づいてモデル化した数理アルゴリズムであり、質のよい解を得るために必要とされる解の探索空間の広さと、実際的な問題である計算時間とのバランスをとりながら準最適解を求めることができるため、組合せ最適化問題に広く利用されている³⁾。

本研究では、ラフ集合による決定表の縮約化において、より大域的な解の探索を可能にする目的で、GA をルール型知識の探索に用いる方法を提案し、この適用可能性についても検証した。

3. 適用例

本研究において、専門家の経験的知識を抽出するために使用した診断事例は、粘板岩を主体とする道路沿いの斜面の崩壊危険度を診断したものである。危険度の診断は高速道路調査会の診断方法⁴⁾に従っている。極小決定アルゴリズムは、ES の知識ベースとして利用する際、新たなデータによる知識ベースの更新や、推論演算にかかる時間を考慮すると、少しでもルール数及び判定要因に対する項目数が少ないほうがよいと考えられる。そのため、ラフ集合論による決定表

キーワード：遺伝的アルゴリズム, 知識獲得, 斜面崩壊, エキスパート・システム, ラフ集合

連絡先：〒569-1095 高槻市霊仙寺町 2-1-1 TEL 0726-90-2402 FAX 0726-90-2491

の縮約化では，不要な判定要因を取り除く際に，残りの判定要因の数が最小となるものを求め，その組合せについてのミール型知識を抽出する方法をとった．これに対し，GA を用いた方法では，決定表中のすべての項目に対して，取り除くか否かを 1 と 0 によって表現し，これを項目数だけ並べてひとつの染色体とした．個体の適応度を算出する評価関数は

$$F = \{X + a(B_2Y - P_1M) + bB_1Z\} \cdot P_2^N \quad (1)$$

とした．ここで， X, Y, Z はそれぞれ，元の表と比較して，取り除かれた項目の数，ルール数，判定要因数である． M, N はそれぞれ，その染色体を元に作成した極小決定アルゴリズムを元の表の事例に適用した際に，矛盾を含む診断となる事例の数，カバーされなくなる事例の数である． B_1, B_2, P_1, P_2 は以上の項目に対する重み付けのための係数である． a, b は定数で，それぞれ，元の表（重複する事例は除く）に含まれる判定要因数と事例数である．

4. 結果と考察

以下にラフ集合論のみに従って得られた極小決定アルゴリズムと，GA を組み合わせる方法によって得られた極小決定アルゴリズムのなかから，ルール数及び判定要因に対する項目数で比較して最小となるものの一例を示す．

表-1 極小決定アルゴリズム

ルール No.	判定要因					危険度
	1	3	4	8	11	
1				2		B
2				3		B
3			1			A
4			3		1	B
5		2			2	B
6		2			3	C
7	2					A
8	3	1			2	A
9	3	1	3			B
10	4				2	B
11	4			4	3	C

表-2 極小決定アルゴリズム(GA)

ルール No.	判定要因							危険度
	1	2	4	8	9	11	12	
1				2				B
2				3				B
3			1					A
4			3			1		B
5		1	2	4				A
6		2				2		B
7	3						2	C
8	3				2			B
9	4			4		3		C

表-1 及び表-2 から，ラフ集合論のみでも，GA を用いた場合でも，最後に残った判定要因の多くが共通しており，これらの判定要因が専門家の特に重視する点であると考えられることができる．また，表 1 及び表-2 において，導かれたルール数と判定要因に対する項目数はそれぞれ 11 ルール・21 項目，9 ルール・17 項目であり，ラフ集合論のみを用いて判定要因の数が最も少なくなる組合せのみから導いた場合よりも，GA を利用してより大きな解空間から導き出した極小決定アルゴリズムのほうが若干小さく，より良い解であるということができる．

5. おわりに

専門家の斜面崩壊危険度診断事例からラフ集合論に基づいて極小決定アルゴリズムを導くための方法として，GA を適用した．その結果，ラフ集合論による組合せの総当たりによる方法と比較して，同等以上の結果が得られることがわかった．また，すべての場合において単純に判定要因を多く取り除けばよいとは限らず，GA によって大域的に解を探索することによって，最適解を効率よく探索できることがわかった．今後の課題として，得られた解に含まれる判定要因が，本当に専門家の診断にとって重視されているものなのかどうかを検証していく必要などが考えられる．

本研究は，平成 11 年度関西大学学部共同研究費によって行った．

参考文献

- 1) 広兼道幸・古田均・中井真司・三雲是宏：斜面の崩壊危険度診断事例からのラフ集合を用いたルール型知識の抽出方法，土木学会論文集，No.582/III-41，pp.285-294，1997.
- 2) Z. Pawlak: Rough Sets-Theoretical Aspects of Reasoning about Data-，Kluwer Acad. Pub，1991.
- 3) 電気学会編：遺伝的アルゴリズムとニューラルネット，コロナ社，pp.1-34，1998.
- 4) 高速道路調査会：地滑り及び斜面崩壊の防止対策の調査手法に関する研究，pp.117-131，1977.