A Granular Computing Approach to Decision Analysis using Rough Set Theory

Iftikhar U. Sikder
Department of Computer and Information Science,
Cleveland State University
Cleveland, OH 44115, USA

Abstract

This paper presents a granular computing approach to decision analysis using rough set theory and its variable precision extension. The multiattribute structure of decision domain is mapped to the notions of equivalence relations of rough set theory. It allows expressing decision categories in terms of approximation space wherein a decision class can be approximated through the partition of boundary region. The variable precision extension of rough set the memberships function is used to generalize the lower and upper approximations. The decision analytic problems can be mapped into rough set theory at predefined precision level.

Keywords
Decision Analysis, Approximate Reasoning, Multi-valued logic, multiple criteria.

1. Introduction

In recent years there is a growing interest in application of granular computing models to decision analysis[1-3]. This trend is triggered by the fact that there is inadequate support to represent and characterize uncertainty when the decision makers or analysts are required to extract knowledge and induce semantic classification[3]. The part of the problem lies in model representation of decision domain in terms of idealized abstraction of primitives in crisp boundary sets. It is necessary to understand how the decision theoretic attributes or evidences suggesting various conceptual or thematic classifications can be understood in a broader approach to allow objective classification by considering the uncertainty intrinsic in decision processes.

2. Rough Set Theory

Rough set theory [4] allows one to characterize a decision class in terms of elementary attribute sets in an approximation space. Decison categories can be represented in the form ($U, C \cup \{d\}$), where d∉C is the decision attribute or the thematic feature and *U* is the closed universe which consists of non-empty finite set of objects (a number of decision categories) and C is a non-empty finite set of attributes that characterizes a decision category such that $c: U \rightarrow V_c$ for every c ∈ C, V_c is a value of attribute c. This is achieved by means of information granulation or indiscernibility is at the heart of rough set theory. A finer granulation means more definable concept. For $P \subseteq C$ the granule of knowledge about a forest with respect to indiscernibility relation can be represented as:

$$Ind(P) = \{x, x') \in U^2 \mid \forall c \in Pc(x) = c(x')\} \quad (1)$$

Thus, the objects x and x' are indiscernible from each other if $(x, x') \in Ind(P)$ and the decision about the presence or absence of a given category is approximated by lower and upper approximation of decision concept as follows:

$$P\underline{X} = \{x \in U \mid Ind(x) \subseteq X \quad (2)$$

$$P\overline{X} = \{x \in U \mid Ind(x) \cap X \neq \emptyset\} \quad (3)$$

Thus, the objects in $P\underline{X}$ can be classified with certainty as the on basis of knowledge while, the objects in $P\overline{X}$ can be only classified as the possible occurrence of decision class and the boundary region ($P\overline{X} - P\underline{X}$) represents the uncertainty in decisive classification.

Since, in real world it is difficult to identify all possible causal attributes, it will be necessary to establish a methodology to identify the critical attributes by eliminating redundant attributes using feature reduction or knowledge compression methods in rough set knowledge systems. This can be achieved by removing the attributes whose removal will not change the indiscernibility relation. A discernibility function $f_A(a_1,.....,a_m) = \wedge\{c_{ij} : 1 \le i < j \le n \,\&\, c_{ij} \ne \emptyset\}$ representing the prime implicants of candidate attributes ($(a_{i1} \wedge .... \wedge a_m)$ generates the minimal set of attributes. While computing prime implicants is an NP-Hard [5] problem, it is possible to use heuristic algorithm (e.g., genetic algorithm [6, 7] or dynamic reducts [8] to generate a computationally efficient set of minimal attributes. Further, it is possible to discover the degree of dependency of the attributes to uncover causal inference of the decision process. The advantage is that the output will provide a mathematically rigorous means to trace or back track the causal links and support transparent decision processes. Thus the rules induced by rough set knowledge discovery process can be regarded as data pattern that represents relationship between multiple decision makers as well as qualitative knowledge. The minimal set of satisfactory rules will provide the means to generate a predictive measure where the associated risk can be evaluated easily. The rules can also provide new insight about the potential impact of some decision Table 1 illustrates an example of sample reading at different locations or granular units of knowledge set.

Table 1: An instance of Decision System in Rough Set

Object ID	[A1]	[A2]	[A3]	Decision
S1	Low	Low	High	A
S2	Low	Low	High	A
S3	High	Low	High	D
S4	Low	High	Low	B
S5	Low	High	Low	C

In this case, we can define the following partitions based of the indiscernibility relations:
IND ({*[A1]*, [*A2*], *[A3]*}) = {*S1*, *S2*}, {*S3*}, {*S4, S5*}}, as shown in table 2.

Table 2: Equivalence Classes

	[A1]	[A2]	[A3]
E1: [*S1, S2]*	Low	Low	High
E2: [*S3*]	High	Low	High
E3: [*S4, S5*]	Low	High	Low

A discernibility matrix defines each equivalence class with respect to one row and column. For a set of attributes in (U, A) the discernibility matrix MT(B) is an n x n matrix such that

$M_T(B) = m_{ij} = |U / IND(B)|$, where

$M_T(i, j) = \{a \in B \mid a(E_i) \ne a(E_j)\}$,for *i, j = 1,2 , ..n*

Table 3: The Discernibility Matrix

	E1	E2	E3
E1	$\emptyset$		
E2	{*[A1]*}	$\emptyset$	
E3	{*[A2], [A3]*}	{ *[A1], [A2], [A3]* }	$\emptyset$

Table 3 shows a symmetric discernibility matrix where each entry in the matrix represents attributes, the value of which render the equivalence classes different.

2.1 Discernibility Function

Using the discernibility matrix, it is possible to calculate the reducts or most important attributes of the information systems. The discernibility function fA for an information system is a Boolean function defined as follows:

$$f_A = \bigwedge_{i,j \in \{1..n\}} \bigvee m_{ij}(E_i, E_j)$$

where $n = |U/IND(B)|$, and a disjunction is taken for all set of Boolean variable corresponding to the element of discernibility matrix.

For example, the discernibility function of the above information system is:

$f_A([A1], [A2], [A3]) = [A1] \cap ([A2] \cup [A3]) \cap ([A1] \cup [A2] \cup [A3])$

The prime implicants of f_A provides the minimal subsets of attributes.

2.2 Reducing Attributes: Reducts

A reduct represents an attribute subset $B \subseteq A$ of an information systems such that after removal of an attribute/s from an equivalence class it preserves indiscernibility relation. In other words, all attributes in a reduct is indispensable. The set of several such minimal attributes are called reducts. The set of prime implicants of the discernibility function determines the reducts. Using the discernibility function, we can determine the reducts of previous example. Each function is then minimized in product form:

$f_A([A1], [A2], [A3]) = [A1] \cap ([A2] \cup [A3]) \cap ([A1] \cup [A2] \cup [A3])$
$f_A([A1], [A2], [A3]) = ([A1] \cap [A2]) \cup ([A1] \cap [A3])$
Therefore, reducts are: *Reduct 1* = {[*A1*], [*A2*]}, *Reduct 2* = {[*A1*], [*A3*]}

2.3 From Reducts to Rules

Rules are generated from reducts in the form: $\alpha \rightarrow \beta$, read as "if α then β" where the antecedent (α) represents the set of conditions or set of conjunctive attributes and corresponding values and the consequent (β) represents the decision class or their disjunction. Transforming reducts to rules involves linking the attribute value of the object from which reducts are generated to the correspondent attributes of the reduct. Although finding equivalence class is computationally straightforward process, computing minimal reducts is NP-hard [9]. Therefore, computing reduct is a non-trivial task. However, there are various approximation algorithms for calculating reduct. These include genetic algorithm for computing minimal hitting sets [10], algorithm based on dynamic reducts, based on greedy approach of set covering heuristics [11], brute force approach [12]. However, once reducts are found, extracting rules from decision table is relatively an easy process. From the previous example, the derived reducts can be used to extract the following rules:

R1: [A1] (Low) AND [A2](Low) => Decision (A)
R2: [A1] (High) AND [A2](Low) => Decision (D)
R3: [A1] (Low) AND [A2](High) => Decision (B) OR Decision (C)
R3: [A1] (Low) AND [A3] (High) => Decision (A)
R4: [A1] (High) AND [A3] (High) => Decision (D)
R5: [A1] (Low) AND [A3] (Low) => Decision (B) OR Decision (C)
Using these rules it is now possible to classify new instances of decision class.

3. Decision Approximation Evaluation by Granularity Measures

The partition induced by the equivalence classes can be mapped with different set of decision classes and to their lower and upper approximations. The set of objects that certainly are members of a class are assigned to the lower approximation region. The upper approximation is a set consisting of objects that are possibly members of a class with respect to the attributes or given knowledge. The difference between the upper and lower approximation is the boundary region or the region of uncertainty. The rough set defined in this sense is based on the concept of total membership function. In a variable precision rough set, [13] the definition of lower and upper approximation is referred with respect to a variable precision where the lower and upper approximations are a special case.

$$\underline{A}_{\pi} X = \{x \in U \mid \mu_A^X(x) \geq 1 - \pi\}$$
$$\overline{A}_{\pi} X = \{x \in U \mid \mu_A^X(x) \geq \pi\}$$

By using rough memberships, the lower and upper approximations can be generalized to arbitrary levels of precision $\pi \in [0,0.5]$. This is based on the concept of the boundary region thinning of variable precision model where we can look at the distribution of decision values within each equivalence class, and exclude those decisions having lower frequency threshold of π. In this case, the lower frequency decision values can be treated as "noise". Further, the concept of sensitivity and specificity can be used to evaluate the approximation [14]. These evaluation parameters are defined as follows:

$$sensitivity(A, \pi, X) = \frac{|\underline{A}_{\pi} X \cap X|}{|X|}$$

$$specificity(A, \pi, X) = \frac{|(U - \overline{A}_{\pi} X) \cap (U - X)|}{|U - X|}$$

Sensitivity indicates the number of objects correctly approximated as members divided by the actual number of members. Specificity is interpreted as the number of object correctly approximated as non-members divide by the actual number of non-members. Accuracy is defined as the ratio of total number of correctly approximated object to the total number of objects. Accuracy is weighting between sensitivity and specificity.

$$accuracy(A, \pi, X) = \frac{|X|}{|U|} sensitivity(A, \pi, X) + \frac{|U - X|}{|U|} specificity(A, \pi, X)$$

4. Conclusions

The application of rough set techniques to multiattribute decision analysis and classification problem offers useful semantics for handling imprecision and uncertainty in finite resolution decision domain. Unlike the classical approach where there is rigorous requirement of decision parameters and conditionings attributes the rough set approach provides flexibility and objective approximation schemes. It has been shown that variable precision memberships function can be used to generalize the lower and upper approximations to arbitrary levels of precision. Thus, the rough set knowledge induction process can be implemented for handling decision analysis problems to incorporate qualitative knowledge in the knowledge induction process.

References

1. Greco, S., B. Matarazzo, and R. Slowinski, *Rough sets theory for multicriteria decision analysis.* European Journal of Operational Research, 2001, 129(1): pp. 1-47.
2. Thangavel, K. and A. Pethalakshmi, *Dimensionality reduction based on rough set theory: A review.* Applied Soft Computing Journal, 2008, 9(1): pp. 1-12.
3. Sikder, I., *Rough set and granular computing in spatial information*, in *Handbook of Research on Geoinformatics*, H.A. Karimi, Editor. 2008, Information Science Reference.
4. Pawlak, Z., *Rough sets: Theoretical aspects of reasoning about data*. System Theory, Knowledge Engineering and Problem Solving, vol. 9. 1991, Kluwer, Dordrecht.
5. Skowron, A. and J. Grzymalla-Busse, *From rough set theory to evidence theory*, in *Advances in the Dempster-Shafer Theory of Evidence*, R. Yager, M. Fedrizzi, and J. Kacprzyk, Editors. 1994, John Wiley & Sons, Inc.: New York. pp. 192-271.
6. Wroblewski, J., *Finding minimal reducts using genetic algorithms*, in *Proceedings of the International Workshop on Rough Sets Soft Computing at Second Annual Joint Conference on Information Sciences (JCIS'95)*, P.P. Wang, Editor. 1995: Wrightsville Beach, NC. pp. 186-189.
7. Wang, X., W. Wan, and X. Yu. *Rough set approximate entropy reducts with order based particle swarm optimization*. in *ACM/SIGEVO Summit on Genetic and Evolutionary Computation*. 2009. Shanghai, China
8. Bazan, J.G., A. Skowron, and P. Synak. *Dynamic Reducts as a Tool for Extracting Laws from Decisions Tables*. in *8th International Symposium on Methodologies for Intelligent Systems*. 1994. London, UK: Springer-Verlag.

9. Skowron, A. and C. Rauszer, *The discernibility matrices and functions in information systems. Intelligent Decision Support*. Handbook of Advances and Applications of the Rough Set Theory, ed. R. Slowinski. 1992, Kluwer, Dordrecht, Netherlands, pp. 331-362.
10. Vinterbo, S. and A. Øhrn, *Minimal approximate hitting sets and rule templates.* International Journal of Approximate Reasoning, 2000, 25(2): pp. 123-143.
11. Johnson, D., *Approximation algorithms for combinatorial problems.* Journal of Computer and System Sciences, 1974, 9: pp. 256-278.
12. Synak, P., *Rough Set Expert System User's Guide: Version 1.0.* 1995, Institute of Mathematics, Warsaw University, Poland.
13. Ziarko, W. *Set Approximation Quality Measures in the Variable Precision Rough Set Model*. in *Soft Computing Systems - Design, Management and Applications, HIS 2002* 2002, Santiago, Chile: IOS Press.
14. Komorowski, J. and A. Øhrn, *Modelling prognostic power of cardiac tests using rough sets.* Artificial Intelligence in Medicine, 1999, 15, pp. 167-191.