Formulating Game Strategies in Game-Theoretic Rough Sets

Nouman Azam and JingTao Yao

Department of Computer Science, University of Regina,
Regina, Saskatchewan, Canada S4S 0A2
{azam200n,jtyao}@cs.uregina.ca

Abstract. The determination of thresholds (α, β) has been considered as a fundamental issue in probabilistic rough sets. The game-theoretic rough set (GTRS) model determines the required thresholds based on a formulated game between different properties related to rough sets approximations and classification. The game strategies in the GTRS model are generally based on an initial threshold configuration that corresponds to the Pawlak model. We study different approaches for formulating strategies by considering different initial conditions. An example game is shown for each case. The selection of a particular approach for a given problem may be based on the quality of data and computing resources at hand. The realization of these approaches in GTRS based methods may bring new insights into effective determination of probabilistic thresholds.

1 Introduction

The probabilistic rough set model has been recognized as a major extension, improvement and generalization of the Pawlak rough set model [10]. The model utilizes a pair of probabilistic (α, β) thresholds to determine the division between probabilistic positive, negative and boundary regions [10]. A fundamental issue in probabilistic rough sets is the computation or determination of the (α, β) threshold parameters [11]. Several attempts have been made recently in this regard including decision-theoretic, game-theoretic, information-theoretic, optimization based and risk based approaches [1, 3, 4, 5, 6, 7]. Despite these attempts, it might still be premature at this point of time to come up with a solution that is universally accepted and convince the majority (if not all) of the audience for its superiority. For now, the need for further research remains in order to obtain more interesting results.

The game-theoretic rough set (GTRS) model has recently provided an alternative way for determining the probabilistic thresholds [4]. It utilizes a game-theoretic environment in determining these thresholds by analyzing and directing towards the optimization of one or more characteristics of the rough set model. Particularly, the thresholds are computed based on a game between different properties related to rough sets based approximation, classification or decision making in order to reach a suitable tradeoff.

P. Lingras et al. (Eds.): RSKT 2013, LNAI 8171, pp. 145–153, 2013.

The strategies in GTRS are generally formulated based on an initial threshold configuration $(\alpha, \beta) = (1, 0)$ which corresponds to the Pawlak model [1, 2]. This only allows for the formulation of strategies in terms of decreasing levels for threshold α and increasing levels for threshold β (considering $0 \leq \beta < \alpha \leq 1$ in the probabilistic rough set model). It seems that the motivation or rationale behind this approach is to obtain a model that is at least better than the Pawlak model based on some considered performance criteria. The approach is useful for configuring the thresholds, it may not necessarily provide an overall better model. We propose various approaches for formulating strategies by considering different initial conditions. A game is implemented for each approach and the threshold modification trend based on a repetitive game is examined. It is hoped that these approaches may further improve and enhance the process of threshold modification and the quality of obtained thresholds.

2 Problem Statement

A main result of probabilistic rough sets is that the rules for determining the three regions are given by,

$$\begin{array}{ll} \text{Positive:} & \text{if } P(C|[x]) \geq \alpha, \\ \text{Negative:} & \text{if } P(C|[x]) \leq \beta, \text{ and} \\ \text{Boundary:} & \text{if } \beta < P(C|[x]) < \alpha. \end{array} \tag{1}$$

where $P(C[x])$ denotes the conditional probability of an object x to be in C given that the object is in $[x]$ and $0 \leq \beta < \alpha \leq 1$. The division between the three regions is based on the probabilistic thresholds (α, β) [9]. The determination and interpretation of thresholds are among the fundamental issues in probabilistic rough sets [11]. There are at least three approaches to determine the thresholds based on decision theory, game theory and information theory that lead us to decision-theoretic rough set (DTRS) [9], game-theoretic rough set (GTRS) [2, 4] and information theoretic rough set (ITRS) [3] models, respectively.

The GTRS model determines the threshold parameters based on a formulated game. A typical game consists of a tuple $\{P, S, u\}$, where:

- P is a finite set of n players, indexed by i,
- $S = S_1 \times ... \times S_n$, where S_i is a finite set of strategies available to player i.
- $u = (u_1, ..., u_n)$ where $u_i : S_i \longmapsto \Re$ is a real-valued utility or payoff function for player i.

The GTRS model considers the players in the form of multiple criteria. Each criterion represents a particular aspect of interest like accuracy or applicability of decision rules. Suitable measures are selected to evaluate these criteria in the context of rough sets based approximation and classification. Each criterion is affected by considering different (α, β) threshold configurations. The strategies are therefore formulated in terms of changes in probabilistic thresholds [1].

The payoff functions represent possible gains, benefits or performance levels achieved by considering different modification in threshold levels.

It is not generally suitable to look into the entire range of threshold values within a single GTRS based game. A repetitive or iterative game is generally used where at each iteration the game outcome is used in directing towards optimal threshold values. In existing GTRS based approaches, the initial (α, β) pair is considered as (1,0) that corresponds to the Pawlak model. We suggest and investigate additional approaches for formulating strategies by considering different initial conditions for determining effective threshold values.

3 Approaches for Formulating Strategies in GTRS

This section introduces four approaches for formulating strategies with GTRS. The game structure and threshold modification trend is discussed for each case.

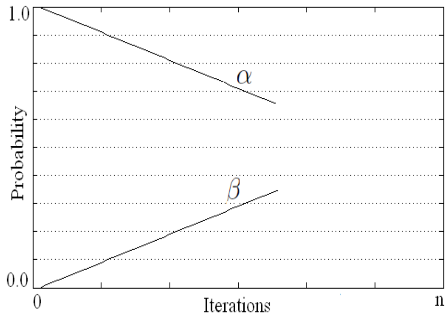
3.1 The Two Ends Approach

The generally used approach for formulating strategies in GTRS is to consider suitable decreasing levels for threshold α and increasing levels for threshold β. Examples of this approach can be found in [1, 2, 4]. The strategies formulated in this way commonly consider an initial configuration of thresholds values, i.e. $(\alpha, \beta) = (1, 0)$ that corresponds to the Pawlak model. We call this approach as the two ends approach since the threshold values are being modified from the two extreme ends.

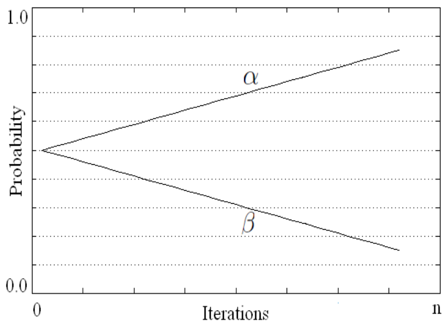
Table 1. Game for two ends approach

		P_2		
		$s_1 = \alpha_\downarrow$	$s_2 = \beta_\uparrow$	$s_3 = \alpha_\downarrow\beta_\uparrow$
P_1	$s_1 = \alpha_\downarrow$	....	.....	.....
	$s_2 = \beta_\uparrow$	....	.....	.....
	$s_3 = \alpha_\downarrow\beta_\uparrow$	....	.....	.....

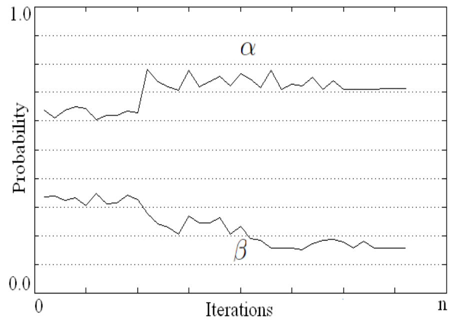
An example game based on this approach is presented in the form of Table 1. Each player in this game considers three strategies, namely $s_1 = \alpha_\downarrow$ (decrease α), $s_2 = \beta_\uparrow$ (increase β), and $s_3 = \alpha_\downarrow\beta_\uparrow$ (decrease α and increase β). The increases or decreases may be set by the user or may be defined in terms of the utilities attained by the players. The outcome of this game may be used to repeat the game based on new values of the thresholds. As the game repeats, the threshold α is continuously decreased while threshold β is increased. The amount of an increase or decrease depends on the outcome of the implemented game.



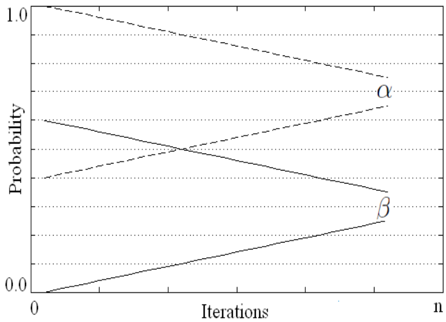
(a) The two ends Approach



(b) The middle approach



(c) The random approach



(d) The range approach

Fig. 1. The four approaches for threshold determination

Figure 1(a) shows the general threshold modification trend with this approach. The modifications in the threshold values are not necessarily linear with respect to iterations. The stop criteria with this approach should be defined to ensure that the process stops before the threshold α becomes less than or equal to β. This approach may be useful when the data are of high quality and the classes or concepts are well defined. A minimum size for the boundary region may be expected in this case. One can make certain decisions with high accuracy rate while keeping the value of α close to 1.0 and β close to 0.0. This means that an effective model may be obtained by considering some minor adjustments to threshold values $(\alpha, \beta) = (1, 0)$.

3.2 The Middle Approach

An alternative approach for formulating strategies is to consider the threshold modification from an initial threshold setting given by $\alpha = \beta$ that corresponds to the two-way decision model. Considering the constraint $\beta < \alpha$, a formulated game based on this approach should consider strategies for increasing α and decreasing β. In some sense this approach can provide an opposite mechanism for

Table 2. Game for middle start approach

		P_2		
		$s_1 = \alpha_{\uparrow}$	$s_2 = \beta_{\downarrow}$	$s_3 = \alpha_{\uparrow}\beta_{\downarrow}$
P_1	$s_1 = \alpha_{\uparrow}$	....	.....	....
	$s_2 = \beta_{\downarrow}$	....	.....	....
	$s_3 = \alpha_{\uparrow}\beta_{\downarrow}$	....	.....	....

Table 3. Game for random start

		P_2	
		$s_1 = \alpha_{\uparrow}$	$s_2 = \alpha_{\downarrow}$
P_1	$s_1 = \beta_{\uparrow}$	....	.....
	$s_2 = \beta_{\downarrow}$	....	.....

threshold configuration as compared to the two ends approach (where threshold α keeps decreasing while β keeps increasing). As the thresholds are being modified from a common or middle value, we name this approach as middle approach.

An example game for this approach may be implemented as shown in Table 2. The strategies may be interpreted as $s_1 = \alpha_{\uparrow}$ (increase α), $s_2 = \beta_{\downarrow}$ (decrease β), and $s_3 = \alpha_{\uparrow}\beta_{\downarrow}$ (increase α and decrease β). When this game is played repeatedly, the threshold α is expected to increase and β is expected to decrease. Figure 1(b) shows the expected development in the two threshold values based on the repeated game. The stop conditions in this approach should be carefully designed such that the iterative process stops before the Pawlak model is reached. This approach may be useful to compare the probabilistic two way decision model and the probabilistic three-way decision model. Particularly, it can provide further insights into the performance related issues associated with the two models.

This approach may be used when the data are of low quality and involve a high level of uncertainty. In such cases we expect many objects in the boundary region leading to its larger size. The number of available certain decisions are very limited. The objective in such situations is to reduce the boundary size to allow for some certain decisions at a cost of some decrease in the level of accuracy. The middle start which starts from zero sized boundary can provide useful configuration of thresholds under these conditions.

3.3 The Random Approach

We may consider a random point for starting the threshold configuration with GTRS. It is assumed that we do not have any knowledge about the modification direction that will provide effective threshold values. In other words, we are not sure wether to increase or decrease a particular threshold. The formulated strategies should therefore provide options for both increasing or decreasing a particular threshold. This means that the strategies will allow us to investigate effective threshold values in the neighborhood of the starting random point.

Table 3 presents an example game for this approach. Here the strategies for the two players are different. Players 1 has the strategies $s_1 = \beta_{\uparrow}$ (increase β)

and $s_2 = \beta_{\downarrow}$ (decrease β) and player 2 has the strategies $s_1 = \alpha_{\uparrow}$ (increase α) and $s_2 = \alpha_{\downarrow}$ (decrease α). Such a game may be realized when player 1 is considering some property of the negative region while player 2 is reflecting the same or some other property of the positive region. Figure 1(c) presents the general threshold modification trend. An implementation of this approach should provide a configuration that is at least better than the initial random point. However, an overall optimal configuration may be not be necessarily achieved. Finally, this approach may be suited to applications that are associated with an intermediate level of uncertainty where the effective threshold values can be located anywhere in the threshold space.

3.4 The Range Approach

The strategies may also be formulated by considering a possible range of values for the thresholds. It may not be feasible to evaluate and consider the entire set of values contained in the range within a single game, however, some selected values from the range may be represented as possible strategies. The game may start from a wider range which is iteratively reduced to a finer range based on a game outcome in a repeated game.

The game in Table 4 may be used to implement this approach. Considering an initial range for threshold α as $[0.5, 1.0]$, the strategies $s_1 = \alpha_1, s_2 = \alpha_2, ..., s_n = \alpha_n$ are representing different values in the considered range. Realizing an order among the strategies such as $\alpha_1 < \alpha_2 ... < \alpha_n$. The strategy α_1 may represent the lower value in the range, i.e. 0.5 and the α_n may represent the upper value in the range, i.e. 1.0. The other strategies may represent intermediate values taken at some specified intervals within the range. Similar interpretation may apply to strategies $s_1 = \beta_1, s_2 = \beta_2, ..., s_n = \beta_n$. The range may be reduced repeatedly by some specified factor, e.g. the range [0.5,1.0] for α may be reduced by a factor of 2 as $(1.0 - 0.5)/2 = 0.25$. The new range may be centered around the threshold values determined by the game outcome. Figure 1(d) presents the general trend in modifying thresholds with this approach. The approach may be useful when we are faced with tight computing constraint and quick convergence or determination of thresholds is desired.

Table 4. Game for range based approach

		P_2		
		$s_1 = \alpha_1$	.....	$s_n = \alpha_n$
P_1	$s_1 = \beta_1$	....	.....	....
	...	....	.....	....
	$s_n = \beta_n$	....	.....	....

4 Threshold Configuration with the Two Ends Approach

We provide an example for the two ends approach which can be used to construct examples for the other approaches. The example is similar to those discussed in [1, 2, 3]. Table 5 represents probabilistic information about a category or concept C based on a partition consisting of 18 equivalence classes. An equivalence class is represented as X_i, and its conditional probability with C as $P(C|X_i)$.

Table 5. Probabilistic information of a concept C

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
$Pr(X_i)$	0.034	0.099	0.132	0.017	0.068	0.017	0.056	0.049	0.049
$Pr(C\|X_i)$	1.0	0.96	0.91	0.86	0.81	0.77	0.71	0.64	0.53
	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}
$Pr(X_i)$	0.115	0.072	0.01	0.119	0.019	0.042	0.009	0.047	0.046
$Pr(C\|X_i)$	0.49	0.43	0.38	0.31	0.27	0.22	0.15	0.09	0.02

Let us consider the game shown in Table 1 for implementing the two ends start approach. Considering the players in the game as the properties of accuracy and generality of the rough set model. For a group containing both positive and negative regions we may define these measures as [2],

$$Accuracy(\alpha,\beta) = \frac{\text{Correctly classified objects by POS}_{(\alpha,\beta)} \text{ and NEG}_{(\alpha,\beta)}}{\text{Total classified objects by POS}_{(\alpha,\beta)} \text{ and NEG}_{(\alpha,\beta)}}, \tag{2}$$

$$Generality(\alpha,\beta) = \frac{\text{Total classified objects by POS}_{(\alpha,\beta)} \text{ and NEG}_{(\alpha,\beta)}}{\text{Number of objects in } U}. \tag{3}$$

where $\text{POS}_{(\alpha,\beta)}$ and $\text{NEG}_{(\alpha,\beta)}$ are the probabilistic positive and negative regions. For each X_i, $X_i \subseteq \text{POS}_{(\alpha,\beta)}$ if $P(C|X_i) \geq \alpha$ and $X_i \subseteq \text{NEG}_{(\alpha,\beta)}$ if $P(C|X_i) \leq \beta$. This means that for $(\alpha,\beta) = (0.9, 0.1)$, we have, $\text{POS}_{(0.9,0.1)} = \bigcup\{X_1, X_2, X_3\}$ and $\text{NEG}_{(0.9,0.1)} = \bigcup\{X_{17}, X_{18}\}$.

Considering U as the total number of objects, number of objects classified by positive and negative regions can be calculated as [3],

$$\text{Classified objects by POS}_{(\alpha,\beta)} = \sum_{P(C|X_i)\geq\alpha} P(X_i) \times U, \text{ and}$$

$$\text{Classified objects by NEG}_{(\alpha,\beta)} = \sum_{P(C|X_i)\leq\beta} P(X_i) \times U. \tag{4}$$

Moreover, the number of correctly classified objects can be determined as [3],

$$\text{Correctly classified by POS}_{(\alpha,\beta)} = \sum_{P(C|X_i)\geq\alpha} P(C|X_i) \times P(X_i) \times U, \text{ and}$$

$$\text{Correctly classified by NEG}_{(\alpha,\beta)} = \sum_{P(C|X_i)\leq\beta} (1 - P(C|X_i)) \times P(X_i) \times U. \tag{5}$$

Table 6. The example game for the two ends approach

		Generality		
		$s_1 = \alpha_\downarrow$	$s_2 = \beta_\uparrow$	$s_3 = \alpha_\downarrow\beta_\uparrow$
Accuracy	$s_1 = \alpha_\downarrow$	(0.941,0.265)	(0.937,0.179)	(0.946,0.311)
	$s_2 = \beta_\uparrow$	(0.973,0.179)	(0.959,0.127)	**(0.959,0.226)**
	$s_3 = \alpha_\downarrow\beta_\uparrow$	(0.946,0.311)	(0.959,0.226)	(0.941,0.358)

For a threshold pair $(\alpha, \beta) = (0.9, 0.1)$, we can calculate the total number of classified objects by $\mathrm{POS}_{(0.9,0.1)}$ as $(P(X_1) + P(X_2) + P(X_3)) \times U = 0.265 \times U$ and the number of classified objects by $\mathrm{NEG}_{(0.9,0.1)} = (P(X_{17}) + P(X_{18})) \times U = 0.093 \times U$. Similarly, the number of correctly classified objects by $\mathrm{POS}_{(0.9,0.1)} = (P(C|X_1) * P(X_1) + P(C|X_2) * P(X_2) + P(C|X_3) * (P(X_3)) \times U = 0.2492 \times U$ and the number of correctly classified objects by $\mathrm{NEG}_{(0.9,0.1)} = ((1 - P(C|X_{17})) * P(X_{17}) + (1 - P(C|X_{18})) * P(X_{18})) \times U = 0.0879 \times U$. Putting these values in Equations (2) - (3), we have

$$Accuracy(0.9, 0.1) = \frac{(0.2492 + 0.0879) \times U}{(0.265 + 0.093) \times U} = \frac{0.3371}{0.358} = 0.941,$$
$$Generality(0.9, 0.1) = \frac{(0.265 + 0.093) \times U}{U} = 0.358. \qquad (6)$$

Focusing the game in Table 1, each player is allowed to choose from one of the following strategies namely $s_1 = \alpha_\downarrow$ (decrease α), $s_2 = \beta_\uparrow$ (increase β), and $s_3 = \alpha_\downarrow\beta_\uparrow$ (decrease α and increase β). Let us consider a decrease or increase of 5%. Each cell in the Table 1 corresponds to a strategy profile. A threshold pair corresponding to a strategy profile is calculated based on two rules, 1) If only one player plays a strategy of modifying a particular threshold, the value will be determined as an increase or decrease suggested by that player, 2) If both the players play the strategies of modifying a particular threshold, the value will be decided as the sum of the two changes.

Considering an initial threshold configuration of $(\alpha, \beta) = (1, 0)$, we may calculate the threshold pairs corresponding to different strategy profiles. For instance the profile $(s_1, s_1) = (\alpha_\downarrow, \alpha_\downarrow) = (0.9, 0.0)$. The corresponding values for the measures accuracy and generality can be calculated as mentioned above. Table 6 shows the resulting game. The pair of values inside a particular cell represents the utilities of the players. The cell with bold values represent the solution of the game determined by the Nash equilibrium [8]. The corresponding threshold values are given by $(\alpha, \beta) = (\beta_\uparrow, \alpha_\downarrow\beta_\uparrow) = (0.95, 0.1)$. The determined values may be used again to implement a game for the next round. Implementing an iterative game in this fashion will result in the modification sequence of $1.0 \rightarrow 0.95 \rightarrow 0.90 \rightarrow 0.85$ for threshold α and $0.0 \rightarrow 0.1 \rightarrow 0.15 \rightarrow 0.25$ for β. It is noted that these threshold modification trends are similar to those shown in Figure 1(a).

5 Conclusion

The game-theoretic rough set model has recently received some attention for determining effective probabilistic thresholds defining the three probabilistic rough set regions. The GTRS implements a game where the strategies are realized as different levels for modifying the thresholds. In this article, we examine additional approaches for formulating strategies based on different initial conditions. The implementation of these approaches is realized by considering example games corresponding to each approach. The iterative threshold modification with these approaches based on a repetitive game is also discussed. It is argued that some of these approaches may be more appropriate when different types of data and applications are considered. A demonstrative example is included to show the usability of the suggested approaches.

Acknowledgements. This work was partially supported by a Discovery Grant from NSERC Canada and the University of Regina FGSR Dean's Scholarship program.

References

[1] Azam, N., Yao, J.T.: Multiple criteria decision analysis with game-theoretic rough sets. In: Li, T., Nguyen, H.S., Wang, G., Grzymala-Busse, J.W., Janicki, R., Hassanien, A.E., Yu, H. (eds.) RSKT 2012. LNCS, vol. 7414, pp. 399–408. Springer, Heidelberg (2012)
[2] Azam, N., Yao, J.T.: Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets. International Journal of Approximate Reasoning (2013), http://dx.doi.org/10.1016/j.ijar.2013.03.015
[3] Deng, X., Yao, Y.Y.: An information-theoretic interpretation of thresholds in probabilistic rough sets. In: Li, T., Nguyen, H.S., Wang, G., Grzymala-Busse, J.W., Janicki, R., Hassanien, A.E., Yu, H. (eds.) RSKT 2012. LNCS, vol. 7414, pp. 369–378. Springer, Heidelberg (2012)
[4] Herbert, J.P., Yao, J.T.: Game-theoretic rough sets. Fundamenta Informaticae 108(3-4), 267–286 (2011)
[5] Jia, X.Y., Tang, Z.M., Liao, W.L., Shang, L.: On an optimization representation of decision-theoretic rough set model. International Journal of Approximate Reasoning (2013), http://dx.doi.org/10.1016/j.ijar.2013.02.010
[6] Li, H.X., Zhou, X.Z.: Risk decision making based on decision-theoretic rough set: A three-way view decision model. International Journal of Computational Intelligence Systems 4(1), 1–11 (2011)
[7] Liu, D., Li, T., Ruan, D.: Probabilistic model criteria with decision-theoretic rough sets. Information Science 181(17), 3709–3722 (2011)
[8] von Neumann, J., Morgenstern, O., Kuhn, H., Rubinstein, A.: Theory of Games and Economic Behavior (Commemorative Edition). Princeton University Press (2007)
[9] Yao, Y.Y.: Decision-theoretic rough set models. In: Yao, J., Lingras, P., Wu, W.-Z., Szczuka, M.S., Cercone, N.J., Ślęzak, D. (eds.) RSKT 2007. LNCS (LNAI), vol. 4481, pp. 1–12. Springer, Heidelberg (2007)
[10] Yao, Y.Y.: Probabilistic rough set approximations. International Journal of Approximate Reasoning 49(2), 255–271 (2008)
[11] Yao, Y.Y.: Two semantic issues in a probabilistic rough set model. Fundamenta Informaticae 108(3-4), 249–265 (2011)