J.T. Yao, J.P. Herbert, A Game-Theoretic Perspective on Rough Set Analysis,
2008 International Forum on Knowledge Technology (IFKT'08), Chongqing,
Journal of Chongqing University of Posts and Telecommunications, Vol. 20, No. 3,
pp291-298, 2008.

A Game-Theoretic Perspective on Rough Set Analysis

JingTao Yao and Joseph P Herbert
(Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada S4S 0A2)

Abstract
Determining the correct threshold values for the probabilistic rough set approaches has been a heated issue among the community. Existing techniques offer no way in guaranteeing that the calculated values optimize the classification ability of the decision rules derived from this configuration. This article will formulate a game theoretic approach to calculating these thresholds to ensure correct approximation region size. Using payoff tables created from approximation measures and modified conditional risk strategies, we provide the user with tolerance levels for their loss functions. Using the tolerance values, new thresholds are calculated to provide correct classification regions. This will aid in determining a set of optimal region threshold values for decision making.

Key words: Game Theory; Rough Set Analysis; Decision-theoretical Rough Sets; Probabilistic Rough Sets;

CLC number: TP18 **Document code:** A **Article ID** : 1673-825X(2008)-03-0291-08

0. Introduction

Game theory [1] is a powerful method for mathematically formulating competition between two or more entities. These entities, or players, aspire to either achieve a dominant position over the other players or collaborate with each other in order to find a position that benefits all. When aiding in data analysis, game theory is useful for exploring the variety of outcomes that result from utilizing different methods or approaches that are designed to solve a problem. In the end, one may determine the course of actions that a technique should undertake to achieve an outcome observed with a game-theoretic formulation.

Game theory may offer a fresh perspective on rough set analysis. Rough sets have been used to aid in conflict analysis [2], a related field to game theory. In Rough Sets [3] and its extensions [4-7], a set within the universe of discourse is approximated. Rough regions are defined with these approximations. One of the goals of improving the classification ability of rough sets is to reduce the boundary region, thus, reduce the amount of classification uncertainty. The decision-theoretic rough set (DTRS) approach [4] to probabilistic rough sets may particularly benefit from some new insights provided by game theory. This approach utilizes the Bayesian decision procedure to calculate classification regions [7]. Loss functions correspond to the risks pertaining to the classification of an object into a particular rough set region. This gives the user a scientific means for linking their risk tolerances with the probabilistic classification ability of rough sets [8].

Classification ability of a rough set analysis system is a measurable characteristic [9]. The decision-theoretic model observes a lower and upper-bound threshold for region classification [10]. A method for measuring the relationship between accuracy or precision and the expected cost of a classification would help in the adoption of these data analysis techniques [11]. The thresholds in the DTRS approach correspond to the probabilities for inclusion into the positive, negative, and boundary regions. These thresholds are calculated through the analysis of loss function relationships. Game theory may allow us to further study the effects of associated risk on these loss functions and provide the user with a suggestion on how to improve the classification ability of the system by either increasing or decreasing their risk tolerances [12].

In this article, we investigate some possible connections between game theory and rough set analysis. We introduce a method for calculating loss tolerance that utilizes game theory to analyze the effects of modifying the classification risk. This also provides an effective means of determining how much a loss function can fluctuate in order to maintain effective classification ability. It is observed that game theory can formulate a link between loss functions and the changes in associated risk with the classification of an object.

1. Game Theory and Decision-Theoretic Rough Sets

We will provide a brief introduction to game theory and the decision-theoretic rough set approach.

1.1 Data Analysis with Game Theory

Game theory [1] has been one of the core subjects of the decision sciences, specializing in the analysis of decision making in an interactive environment. It arose from the result of trying to mathematically express a simple game, including rules and actions a player of that game would perform.

Many applications or problems can be expressed as a game between two or more players. If a problem or application can be expressed as a game, it can be expressed in a way that some aspects of game theory can be utilized. Therefore, the study of game theory can be thought of as an advanced problem solving technique that can be used in many domains. The disciplines utilizing game theory include those of economics [13, 14], machine learning [15], networking [16], and cryptography [17].

The basic assumption of game theory in terms of usage is that all participating players are rational in terms of attempting to maximize their expected payoffs. This presents problems when compared with neoclassical economics. It narrows the range of possibilities that a party can choose from. Rational behavior is much more predictable than irrational behavior, as opposing parties are able to determine other party's strategies on the basis that they will not do anything that makes their situation worse than before.

In a simple game put into formulation consists of a set of players $O = \{o_1, \ldots, o_n\}$, a set of actions $S = \{a_1, \ldots, a_m\}$ for each player, and the respective payoff functions for each action $F = \{\mu_1, \ldots, \mu_m\}$. Each player chooses actions from S to be performed according to expected payoff from F, usually some a_i maximizing payoff $\mu_i(a_i)$ while minimizing other player's payoff.

Let us look at a classical example of the use of game theory: the prisoners' dilemma. In this game, there are two players, $O = \{o_1, o_2\}$. Each player has been captured by authorities in regards to a suspected burglary. While being interrogated by the police, the prisoners each have a choice of two actions they can perform: to confess and implicate the other prisoner for the crime, or not to confess to the burglary, i.e. $S = \{a_1, a_2\}$, where $a_1 = \text{confess}$ and $a_2 = \text{don't confess}$. The payoff functions for each action correspond to the amount of jail time that prisoner will receive. These payoffs are expressed in Table 1, called a payoff table.

Table 1.Payoff table for the Prisoners' Dilemma.

		o_2	
		confess	**don't confess**
o_1	**confess**	o_1 serves 10 years, o_2 serves 10 years	o_1 serves 0 years, o_2 serves 20 years
	don't confess	o_1 serves 20 years, o_2 serves 0 years	o_1 serves 1 year, o_2 serves 1 year

If prisoner o_1 confesses to the burglary and implicates the other, he will serve either a maximum of 10 years in jail or serve nothing, depending on whether or not prisoner o_2 confesses or not. If o_1 doesn't confess, he will serve at least 1 year in jail or a maximum of 20 years. If both confess, they each serve 10 years. If neither confesses, they both serve 1 year. Clearly, without knowing the other's action, a rational player would choose to confess, as it provides the smallest maximum jail term as well as the smallest overall term.

The above example demonstrates one of the strengths of game theory for aiding data analysis. It provides clarity to complex scenarios where multiple actions influence the outcome in a predictable manner. The decision-theoretic rough set approach to data analysis is one such method where classification ability is configurable by observing different values of conditional risk associated with an action.

1.2 The Decision-Theoretic Rough Set Model

The decision theoretic approach is a robust extension of rough sets for two reasons. First, it calculates approximation parameters by obtaining easily understandable notions of risk or loss from the user [10]. It allows for simpler user involvement instead of having parameters being arbitrarily provided. This is important when users are not qualified to set the parameters and just wish to perform analysis. Second, many application domains could make use of cost or risk annotations. We present a slightly reformulated decision theoretic rough set model in this section, as reported in [4, 7].

Let $P(w_j|\mathbf{x})$ be the conditional probability of an object x being in state w_j given the object description $\mathbf{x}$. The set of actions is given by $\mathcal{A} = \{a_P, a_N, a_B\}$, where a_P, a_N, and a_B represent the three actions to classify an object into $POS(A)$, $NEG(A)$, and $BND(A)$ respectively. Let $\lambda(a_\diamond|A)$ denote the loss incurred for taking action $a_\diamond$ when an object is in A, and let $\lambda(a_\diamond|A^c)$ denote the loss incurred by taking the same action when the object belongs to A^c. This can be given as loss functions $\lambda_{\diamond P} = \lambda(a_\diamond|A)$, $\lambda_{\diamond N} = \lambda(a_\diamond|A^c)$, and $\diamond = P,\ N$, or B.

The expected loss $R(a_\diamond|[x])$ associated with taking the individual actions can be expressed as:

$$
\begin{aligned}
R_P = R(a_P|[x]) &= \lambda_{PP}P(A|[x]) + \lambda_{PN}P(A^c|[x]),\\
R_N = R(a_N|[x]) &= \lambda_{NP}P(A|[x]) + \lambda_{NN}P(A^c|[x]),\\
R_B = R(a_B|[x]) &= \lambda_{BP}P(A|[x]) + \lambda_{BN}P(A^c|[x]),
\end{aligned} \tag{1}
$$

where $\lambda_{\diamond P} = \lambda(a_\diamond|A)$, $\lambda_{\diamond N} = \lambda(a_\diamond|A^c)$, and $\diamond = P,\ N$, or B. R_P, R_N, and R_B are the expected losses of classifying an object into the positive region, negative region, and boundary region respectively.

If we consider the loss function inequalities $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$, we can formulate decision rules based on this division of the universe. The corresponding inequalities $\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$ can further tell us how the universe is divided. We can formulate the following decision rules (PP, NP, BP) based on the set of inequalities above [18]:

$$
\begin{aligned}
&\text{(PP)} \quad \text{If } P(A|[x]) \geq \gamma \text{ and } P(A|[x]) \geq \alpha, \quad \text{decide} \quad POS(A),\\
&\text{(NP)} \quad \text{If } P(A|[x]) \leq \beta \text{ and } P(A|[x]) \leq \gamma, \quad \text{decide} \quad NEG(A),\\
&\text{(BP)} \quad \text{If } P(A|[x]) \geq \beta \text{ and } P(A|[x]) \leq \alpha, \quad \text{decide} \quad BND(A),
\end{aligned}
$$

where,

$$
\begin{aligned}
\alpha &= \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{BP} - \lambda_{BN}) - (\lambda_{PP} - \lambda_{PN})},\\
\gamma &= \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{NP} - \lambda_{NN}) - (\lambda_{PP} - \lambda_{PN})},\\
\beta &= \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{NP} - \lambda_{NN}) - (\lambda_{BP} - \lambda_{BN})}.
\end{aligned} \tag{2}
$$

The Bayesian decision procedure leads to the following minimum risk decision rules (PN - BN) [7]:

$$
\begin{aligned}
&\text{(PN)} \quad \text{If } R_P \leq R_N \text{ and } R_P \leq R_B, \quad \text{decide} \quad \text{POS}(A);\\
&\text{(NN)} \quad \text{If } R_N \leq R_P \text{ and } R_N \leq R_B, \quad \text{decide} \quad \text{NEG}(A);\\
&\text{(BN)} \quad \text{If } R_B \leq R_P \text{ and } R_B \leq R_N, \quad \text{decide} \quad \text{BND}(A);
\end{aligned}
$$

The α, β, and γ parameters define our regions, giving us an associated risk for classifying an object. The α parameter can be considered the division point between the POS region and BND region. Likewise, the β parameter is the division point between the BND region and the NEG region.

These minimum risk decision rules offer us a foundation in which to classify objects into approximation regions. They give us the ability to not only collect decision rules from data frequent in many rough set applications [19], but also the calculated risk that is involved when discovering (or acting upon) those rules.

2. Rough Set Analysis from a Game Theory Perspective

We stated previously that the user could make use of a method of linking their notions of cost (risk) in taking a certain action and classification ability of the classification system. This relationship is possible by analyzing the consequences of each fluctuation of expected cost. Game theory is a powerful mathematical paradigm for analyzing these relationships and also provides methods for achieving optimal configurations for classification strategies.

Game theory could provide a means for the user to change their beliefs regarding the types of decisions they can make [20]. If the classification system is not precise enough, they would not have to change the probabilities on their own. This is beneficial as many users cannot intuitively describe their decision needs in terms of probabilities.

Furthermore, when asked if they can modify their cost beliefs (loss functions), they can perhaps be more successful in this description. We present a five step process for using game theory to aid in rough set analysis in this section. These steps take into account the formulations required in order to observe rough set data analysis as a competition between measures, how to measure the outcomes of strategies, how to view the competition in an organized manner, and how to interpret the results one can achieve using these steps.

2.1 The Game Theory Formulation Process

When using game theory to aid in rough set analysis, there are five procedures to be utilized:

Step 1. Game Formulation
- The game formulation defines what the game contains: the players and what they represent as well as what their overall goals consist of.

Step 2. Strategy Formulation
- The strategy formulation defines the possible actions that the player can undertake.

Step 3. Payoff Measurement
- The payoff measurement defines how the game will measure the effectiveness of the actions defined in Step 2.

Step 4. Competition Implementation
- The competition implementation allows for the observation of the game by collecting the information into payoff tables and examining the relationships between the actions undertaken and the payoffs associated with those actions.

Step 5. Result Acquisition
- The result acquisition interprets the results of the competition.

We will present details of these steps in the following subsections, using the decision-theoretic rough set approach.

2.2 Game Formulation

The first step of the process is to formulate a game. When using game theory to help determine suitable loss functions, we need to correctly formulate the following: a set of players, a set of strategies for each player, and a system of payoff functions. Game theory uses these formulations to find optimal an optimal strategy for a single player or the entire group of players if cooperation (coordination) is wanted. A single game is defined as,

$$G = \{O, S, F\} \tag{3}$$

where G is a game consisting of a set of players O using strategies in S. These strategies are measured using individual payoff functions in F.

To begin, the set of players should reflect the overall purpose of the competition. In a typical example, a player can be a person who wants to achieve certain goals. For simplicity, we will be using competition between two players. With improved classification ability as the competition goal, each player can represent a certain measure, i.e., a set of players,

$$O = \{\phi, \psi\} \tag{4}$$

where ϕ and ψ are measures pertaining to different characteristics of the system. In this case, ϕ represents approximation accuracy and ψ represents approximation precision. Through competition, optimal values are attempting to appear for each measure. Hence, an optimal cooperative solution would have each measure achieving maximum payoff. Although we are measuring accuracy and precision, the choice of measures is ultimately up to the user to decide. We wish to analyze the amount of movement or compromise loss functions can have to when attempting to achieve optimal values for these two measures.

Each measure is effectively competing with the other to win the "game". Here, the game is to improve classification ability. To compete, each measure in O has a set of strategies it can employ to achieve payoff. Payoff is the measurable result of actions performed using the strategies. These strategies are executed by the player in order to better their position in the future, i.e. maximize payoff. Individual strategies, when performed, are called actions. It follows,

$$S_i = \{a_1, \ldots, a_m\} \tag{5}$$

where S_i is the strategy set for a measure $p_i \in O$ and a_j is a j-th action in the strategy set. A total of m actions can be performed for this player. This strategy set must contain actions that are related to the goal of the player. For example, the measure ϕ (representing the first player) measures approximation accuracy in the classification system. The strategy for this particular player would be along the lines of "acquire a maximal value for approximation accuracy as possible". Likewise, the strategy for ψ would be to "acquire a maximal value for approximation precision as possible". The actions in each strategy set, when performed, should be able to fulfill the player's goals of finding optimal values.

Approximation accuracy (ϕ), is the ratio measured between the size of the lower approximation of a set A to the upper approximation of a set A. A large value of ϕ indicates that we have a small boundary region.

To illustrate the change in approximation accuracy, suppose we have player ϕ taking two turns in the competition. For the first turn, it executes action a_1 from it's strategy set. When it is time to perform another turn, the player executes action a_2. Ultimately, since the player's goal is to increase approximation accuracy, we should measure that $\phi_{a_1} \leq \phi_{a_2}$. If this is not the case ($\phi_{a_1} > \phi_{a_2}$), the player has chosen a poor second action from it's strategy set.

The second player, approximation precision (ψ), observes the relationship between the upper approximation and a set. In order to increase precision, we need to make $|\underline{apr}(A)|$ as large as possible. For non-deterministic approximations, Yao [7] suggested an alternative precision measure.

In general, the two measures measure the impact that the loss functions have on the classification ability of the DTRS model. Modifying the loss functions contribute to a change in risk (expected cost). Determining how to modify the loss functions to achieve different classification abilities requires a set of risk modification strategies.

2.3 Strategy Formulation

The second step is to formulate strategies for each player in the game. We wish to emphasize the relationship between the condition risk and the loss functions. In order to increase accuracy, we need to make $|\underline{apr}(A)|$ as large as possible while maintaining the size of $|\overline{apr}(A)|$. Recalling rules (PN, NN, BN), we see that in order to increase the size of the lower approximation, we need decrease the expected loss R_P. This results in more objects being classified into the positive region. An increase R_N and R_B is also desired. This is intuitive when considering that in order for more objects to be classified into $POS(A)$, we need to lower the risk involved in classifying an object into this region.

We see that in order to decrease the value of R_P, we need to decrease one or both of the loss functions λ_{PP} and λ_{PN} (Equation 1). Likewise, to increase R_N, we need to increase either λ_{NP} or λ_{NN}. Finally, to increase R_B, we need to increase λ_{BP} or λ_{BN}. This is summarized in Table 2.

Table 2. The strategy scenario of increasing approximation accuracy.

Action (Strategy)	Goal	Method	Result
a_1 ($-R_P$)	Decrease R_P	Decrease λ_{PP} or λ_{PN}	Larger POS region
a_2 ($+R_N$)	Increase R_N	Increase λ_{NP} or λ_{NN}	Smaller NEG region
a_3 ($+R_B$)	Increase R_B	Increase λ_{BP} or λ_{BN}	Smaller BND region

For the second player, ψ, we need to increase approximation precision. For the deterministic case, in order to increase precision, we need to make $|\underline{apr}(A)|$ as large as possible. Again, recalling rules (PN, NN, BN), we see that in order to increase the size of the lower approximation, we need to decrease the expected loss R_P and to increase R_N and R_B. It has the same strategy set as the first player because we wish to increase the size of the lower approximation. To do this, we need to decrease the risk of classifying an object into the positive region.

In order to increase precision in the non-deterministic case, we need to make $|\overline{apr}(A)|$ as small as possible. Recalling rules (PN, NN, BN), we see that in order to decrease the size of the upper approximation, we need to decrease the expected loss R_N and to increase R_P and R_B. This effectively makes classifying objects into the negative region a lower risk endeavour.

To decrease R_N, we decrease the loss functions λ_{NP} and λ_{NN}. Likewise, to increase R_P, we increase either λ_{PP} or λ_{PN}. Finally, to increase R_B, we need to increase λ_{BP} or λ_{BN}. This strategy set is summarized in Table 3.

Table 3. The strategy scenario for non-deterministic approximation accuracy.

Action (Strategy)	**Goal**	**Method**	**Result**
a_1 ($-R_N$)	Decrease R_N	Decrease λ_{NP} or λ_{NN}	Larger NEG region
a_2 ($+R_P$)	Increase R_P	Increase λ_{PP} or λ_{PN}	Smaller POS region
a_3 ($+R_B$)	Increase R_B	Increase λ_{BP} or λ_{BN}	Smaller BND region

2.4 Payoff Measurement

The third step is to define the payoff functions that measure the effectiveness of the actions performed by each player. Payoff, or utility, results from a player performing an action. For a particular payoff for player i performing action a_j, the utility is defined as the following,

$$\mu_{i,j} = \mu(a_j). \tag{6}$$

A set of payoff functions F contains all μ functions acting within the game G. In this system of accuracy and precision, $F = \{\mu_\phi, \mu_\psi\}$, showing payoff functions that measure the increase in accuracy and precision. A formulated game typically has a set of payoffs for each player,

$$P_i = \{\mu_{i,1}, \ldots, \mu_{i,m}\}, \tag{7}$$

where the player i has m total actions in their strategy set. In this article, we will define the payoffs within the competition table (discussed in the next section) for ϕ as $\phi_{1,j}$ and the ψ payoffs as $\psi_{2,k}$. In our approach, given a strategy set S containing three strategies, the payoffs for ϕ and ψ are as follows,

$$\begin{aligned} P_1 &= \{\phi_{1,1}, \phi_{1,2}, \phi_{1,3}\}, \\ P_2 &= \{\psi_{2,1}, \psi_{2,2}, \psi_{2,3}\}, \end{aligned} \tag{8}$$

reflecting payoffs from the result of the three actions pertaining to the three strategies undertaken by each player, i.e., $\mu_\phi(a_j) = \phi_{1,j}$. This is a simple approach that can be expanded to reflect true causal utility based on the opposing player's actions. This means that not only is an action's payoff dependant on the player's action, but also which strategy the opposing player has chosen.

After modifying the respective loss functions, the function μ_ϕ calculates the payoff via approximation accuracy. Likewise, the payoff function μ_ψ calculates the payoff with approximation precision for deterministic approximations. More elaborate payoff functions could be used to measure the state of a game G, including entropy or other meaningful measures according to the player's overall goals [11].

The payoff functions imply that there are relationships between the measures selected as players, the actions they perform, and the probabilities used for region classification. These properties can be used to formulate guidelines regarding the amount of flexibility the user's loss function can have to maintain a certain level of consistency in the data analysis. As we see in the next section, the payoffs are organized into a payoff table in order to perform analysis.

2.5 Competition Implementation

The fourth step is to express the game consisting of players, strategies, and payoffs in a payoff table structure. To find optimal solutions to either ϕ or ψ (or both), we organize payoffs with the corresponding actions that are performed. Since we are considering both deterministic and non-deterministic approximations, we must construct two payoff tables. The first, for deterministic approximations, is shown in Table 4, and will be the focus of our attention. For non-deterministic approximations, a new strategy set would be required for precision.

The actions belonging to ϕ are shown row-wise, whereas the strategy set belonging to ψ are shown column-wise. In Table 4, the strategy set S_1 for ϕ contains three strategies $S_1 = \{-R_P, +R_N, +R_B\}$ pertaining to actions resulting in a decrease in expected cost for classifying an object into the positive region and an increase in expected cost for classifying objects into the negative and boundary regions. The strategy set for ψ contains the same actions for the second player.

Each cell in the table has a payoff pair $< \phi_{1,i}, \psi_{2,j} >$. In this system, a total of 9 payoff pairs are calculated. For example, the payoff pair $< \phi_{3,1}, \psi_{3,1} >$ containing payoffs $\phi_{3,1}$ and $\psi_{3,1}$ correspond to modifying loss functions to increase the risk associated with classifying an object into the boundary region and to decrease the expected cost associated with classifying an object into the positive region. Measures pertaining to accuracy and precision after the resulting actions are performed for all 9 cases. These payoff calculations populate the table with payoffs so that equilibrium analysis can be performed.

Table 4.Payoff table for ϕ, ψ payoff calculation (deterministic).

		ψ_D		
		$-R_P$	$+R_N$	$+R_B$
ϕ	$-R_P$	$< \phi_{1,1}, \psi_{1,1} >$	$< \phi_{1,2}, \psi_{2,1} >$	$< \phi_{1,3}, \psi_{3,1} >$
	$+R_N$	$< \phi_{2,1}, \psi_{1,2} >$	$< \phi_{2,2}, \psi_{2,2} >$	$< \phi_{2,3}, \psi_{3,2} >$
	$+R_B$	$< \phi_{3,1}, \psi_{1,3} >$	$< \phi_{3,2}, \psi_{2,3} >$	$< \phi_{3,3}, \psi_{3,3} >$

In order to find optimal solutions for either accuracy or precision, we determine whether there is equilibrium within the payoff table. This intuitively means that both players attempt to maximize their payoffs given the other player's chosen action, and once found, cannot rationally increase this payoff.

It follows that if an equilibrium is found within the table, i.e., one or more payoff pairs $< \phi^*_{1,i}, \psi^*_{2,j} >$ are calculated, where for any action a_k where $k \neq i, j$, $\phi^*_{1,i} \geq \phi_{1,k}$ and $\phi^*_{2,j} \geq \phi_{2,k}$ is a optimal solution for determining loss functions. Thus, once an optimal payoff pair is found, the user of the system is provided with the following information: a suggested tolerance level for the loss functions and the amount of change in accuracy and precision resulting from the changed loss functions. Equilibrium is a solution to the amount of change loss functions can undergo to achieve levels of accuracy and precision noted by the payoffs.

2.6 Results Acquisition

The final step is to interpret the results from the observations made on the execution of the game. Observed from decision rules (PN, NN, BN), we can calculate how much the loss functions need to be modified to acquire a certain level of accuracy or precision. There is a limit to amount of change allowable for loss functions. For example, the action of reducing the expected cost R_P. We can reduce this cost any amount and rule (PN) will be satisfied. However, the rules (NN) and (BN) are also affected by the modified R_P, denoted R^*_P. R^*_P must satisfy $R^*_P \geq (R_N - R_P)$ and $R^*_P \geq (R_B - R_P)$. This results in an allowable change of t_{PP} to λ_{PP} and t_{PN} to λ_{PN}.

Assuming that $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$ and $\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$, we calculate the following,

$$t^{max}_{PP} = \frac{\lambda_{BP} - \lambda_{PP}}{\lambda_{PP}}, \quad t^{min}_{PN} = \frac{\lambda_{PN} - \lambda_{BN}}{\lambda_{PN}}. \tag{9}$$

That is, t^{max}_{PP} is the *tolerance* that loss function λ_{PP} can have (t^{min}_{PN} for λ_{PN}). Tolerance values indicate how much change a user can have to their risk beliefs (loss functions) in order to maintain accuracy and precision measures of

$< \phi^*_{1,i}, \psi^*_{2,j} >$. In brief, when selecting a strategy, i.e. ($+R_P$), the game calculates payoffs by measuring the approximation accuracy and prediction that result from modifying the loss functions λ_{PP} and λ_{PN}. The new loss functions, λ^*_{PP} and λ^*_{PN} are used to calculate a new expected loss R^*_P. In order to maintain the levels of accuracy and prediction stated in the payoffs, the user must have new loss functions within the levels of t^{max}_{PP} for λ_{PP} and t^{min}_{PN} for λ_{PN}.

2.7 An Example

Let the following be a series of loss functions for correct classifications, boundary classifications, and incorrect classifications respectively:

$$\lambda_{PP} = \lambda_{NN} = 4, \qquad \lambda_{BP} = \lambda_{BN} = 6, \qquad \lambda_{PN} = \lambda_{NP} = 8. \tag{10}$$

The inequality restrictions for the loss functions hold, i.e., $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$ and $\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$. The cost of a correct classification (λ_{PP} and λ_{NN}) is less then the cost for classifying an object into the boundary region (λ_{BP} and λ_{BN}) and both are strictly less than the cost of an incorrect classification (λ_{PN} and λ_{NP}).

Using Equation (9), we calculate the following:

$$t^{max}_{PP} \quad = \quad \frac{6-4}{4} = \frac{1}{2}\ , \quad t^{min}_{PN} = \frac{8-6}{8} = \frac{1}{4}.$$

Thus, $t^{max}_{PP} = 0.5$ and $t^{min}_{PN} = 0.25$. This means that we can increase the loss function λ_{PP} by 50% and increase the loss function λ_{PN} by 25% and maintain the same classification ability.

These tolerance values can now be used to modify the expected loss strategies of the game. Recalling the strategies in Table 2, we use the tolerance values to calculate how much we can decrease R_P and increase R_N and R_B in order to increase approximation accuracy. Likewise, from Table 3, we can use the tolerance value to calculate how much to decrease R_N and increase R_P and R_B.

3. **Conclusions**

This article provides some insights regarding the use of game theory to aid the rough set analysis process. Specifically, we provide a preliminary study on using game theory for determining the relationships between loss tolerance, expected loss, and threshold value modification in the decision-theoretic approach. We achieve this by formulating the process of loss function modification as goals to be achieved in a competitive environment. By choosing measures of approximation accuracy and approximation precision as players in a game, with individual goals of maximizing their values, we set up a set of strategies that each can perform.

In the game-theoretic approach, we investigate the possibility that a change in values for expected losses for classifying objects can be thought of as actions being performed by a player in a game. The strategies involve decreasing or increasing the user-provided loss functions for classifying objects into rough set regions. Taking into account the actions that are used to modify the loss functions to achieve new accuracy and precision measurements, we can indicate how much a loss function can be modified. This is useful for the users as determining the amount of tolerance they should have when modifying loss functions (risk tolerance) is difficult.

Analyzing the payoff tables indicates a relationship between the two measures and the actions performed to reach those values. By undertaking these risk modification actions, new values of the thresholds can be calculated to reflect balanced classification ability. We conclude that game theory can be a powerful method for governing rough set analysis when adjustable criteria, such as loss functions, are used to influence classification ability.

4. **References**

[1] NEUMANN J V, MORGENSTERN O. Theory of games and economic behavior [M]. Princeton: Princeton University Press, 1944.

[2] SKOWRON A, RAMANNA S, PETERS J F. Conflict analysis and information systems: a rough set approach: RSKT'06: proceedings of Rough Sets and Knowledge Technology, Chongqing, July 24-26, 2006 [C]. LNAI 4062, Berlin, Springer: 233-240.
[3] PAWLAK Z. Rough sets [J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356.
[4] YAO Y Y, WONG S K M. A decision theoretic framework for approximating concepts [J]. International Journal of Man-machine Studies, 1992, 37: 793-809.
[5] ZIARKO W. Variable precision rough set model [J]. Journal of Computer and System Sciences, 1993, 46: 39-59.
[6] GRECO S, MATARAZZO B, SLOWINSKI R. Rough set theory for multicriteria decision analysis [J], European Journal of Operational Research, 2001, 129: 1-47.
[7] YAO Y Y. Decision-theoretic rough set models: RSKT'07: proceedings of Rough Sets and Knowledge Technology, Toronto, May 14-16, 2007 [C]. LNAI 4481, Berlin, Springer: 1-12.
[8] YAO J T, HERBERT J P. Web-based support systems with rough set analysis: RSEISP'07: proceedings of Rough Sets and Emerging Intelligent Systems Paradigms, Warsaw, June 28-30, 2007 [C]. LNAI 4585, Berlin, Springer: 360-370.
[9] GEDIGA G, DUNTSCH I. Rough approximation quality revisited [J]. Artificial Intelligence, 2001, 132: 219-234
[10] YAO Y Y. Probabilistic approaches to rough sets [J]. Expert Systems, 2003, 20: 287-297.
[11] DUNTSCH I, GEDIGA G. Uncertainty measures of rough set prediction [J]. Artificial Intelligence, 1998, 106: 109-137.
[12] HERBERT J P, YAO J T. Game-theoretic risk analysis in decision-theoretic rough sets: RSKT'08: proceedings of Rough Sets and Knowledge Technology, Chengdu, May 17-19, 2008 [C]. LNAI 5009, Berlin, Springer: 132-139.
[13] NASH J. The bargaining problem [J]. Econometrica, 1950, 18: 155-162.
[14] ROTH A. The evolution of the labor market for medical interns and residents: a case study in game theory [J]. Political Economy, 1984, 92: 991-1016.
[15] HERBERT J, YAO J T. A game-theoretic approach to competitive learning in self-organizing maps: ICNC'05: proceedings of the first International Conference on Natural Computation, Changsha, August 27-29, 2005 [C]. LNCS 3610, Berlin, Springer: 129-138.
[16] BELL M G F. The use of game theory to measure the vulnerability of stochastic networks [J]. IEEE Transactions on Reliability, 2003, 52: 63-68.
[17] FISCHER J, WRIGHT R N. An application of game-theoretic techniques to cryptography [J]. Discrete Mathematics and Theoretical Computer Science, 1993, 13: 99-118.
[18] YAO Y Y, ZHAO Y. Attribute reduction in decision-theoretic rough set models [J]. Information Sciences, 2008, To Appear.
[19] HERBERT J, YAO J T. Time-series data analysis with rough sets: CIEF'05: proceedings of Computational Intelligence in Economics and Finance, Salt Lake City, July 21-26, 2005 [C]. 908-911.
[20] HERBERT J P, YAO J T. Rough set model selection for practical decision making: FSKD'07: Proceeding of Fuzzy Systems and Knowledge Discovery, Hainan, August 24-27, 2007 [C]. 203-207.

5. Biographies:

Dr. JingTao Yao is an Associated Professor of Computer Science at the University of Regina, Canada. He received his PhD degree at the National University of Singapore. His research interests include soft computing, data mining, neural networks, rough sets, computational finance, forecasting, electronic commerce, Web intelligence, and Web-based support systems. He can be reached at jtyao@cs.uregina.ca

Joseph P. Herbert is a Ph.D. Candidate in the Department of Computer Science at the University of Regina, Canada. His research interests include rough sets, granular computing, self-organizing maps, machine learning, and game theory. He can be reached at herbertj@cs.uregina.ca