Near Tolerance Rough Sets

Sheela Ramanna

Department of Applied Computer Science
University of Winnipeg
Winnipeg, Manitoba R3B 2E9, Canada
s.ramanna@uwinnipeg.ca

Abstract
This article considers the problem of how to formulate a framework for the study of the nearness of tolerance rough sets (TRS). The solution to the problem stems from recent work on near sets and approach spaces as well as from the realisation that disjoint TRSs can be viewed in the context of approach merotopic spaces. A set of TRSs equipped with a distance function satisfying certain conditions is an example of an approach space. An application of TRS-based approach spaces is given in terms of content-based image retrieval.

Mathematics Subject Classification: 54D35, 54A20, 54E99, 18B30

Keywords Approach space, merotopy, near sets, tolerance rough set

1 Introduction
The problem considered in this paper is how to formulate a framework for the study of the nearness of tolerance rough sets (TRS). TRS theory was introduced in 1994 [13]. The solution to the problem stems from recent work on near sets [19, 18, 20, 21, 28] and from the realisation that disjoint TRSs can be viewed in the context of approach spaces [4, 9, 10, 11, 12, 26, 22, 25], especially merotopic spaces [7, 4, 5, 6, 26, 22, 23]. The basic approach is to consider a nonempty set $X$ of TRSs equipped with a distance function $\rho : \mathcal{P}X \times \mathcal{P}X \rightarrow [0, \infty]$ satisfying certain conditions. In that case, $(X, \rho)$ is an approach space. A collection $\mathcal{A}, \mathcal{B} \in \mathcal{P}^2 X$ is near when $\nu_\rho(\mathcal{A}, \mathcal{B}) = \inf_{B \in \mathcal{B}} \sup_{A \in \mathcal{A}} \rho(B, A) = 0$.

2 Tolerance Rough Sets
A nonempty set $X$ is a rough set if, and only if the approximation boundary of $X$ is nonempty. Rather than the usual partition of $X$ with the indiscernibility relation introduced by Z. Pawlak [17], set approximation is viewed in the
context of a cover of $X$ defined by an tolerance relation $\tau_\Phi$, where $\Phi$ is a set of attributes of $x \in X$ [13]. In this work, let $\Phi = \{\phi_1, \ldots, \phi_i, \ldots, \phi_n\}$ denote a set of **probe functions** representing object features, where $\phi_i : X \rightarrow \mathbb{R}$. Feature vectors (vectors of numbers represented as feature values extracted from objects) provide a basis for set descriptions (see, e.g., [19, 18, 15, 16]). Let $A \subseteq X$ and let $\tau_\Phi(x, y)$ denote a tolerance class (maximal preclass) containing the set of all objects $x \in A$, $y \in X$ such $x \tau_\Phi y$. The upper approximation of $A$ is denoted by $\Phi^*A$, the set of all such tolerance classes that have a nonempty intersection with $A$, i.e. The lower approximation of $A$ (denoted $\Phi^*_A$) is the set of all tolerance classes that are proper subsets in $A$.

**3 Approach Spaces**

The collection of all subsets of a nonempty set $X$ is denoted $\mathcal{P}X = 2^X$ (power set). An **approach space** is a nonempty set $X$ equipped with a distance function $\rho : \mathcal{P}X \times \mathcal{P}X :\rightarrow [0, \infty]$ if, and only if, for all nonempty subsets $A, B, C \in \mathcal{P}X$, $\rho$ satisfies conditions (A.1)-(A.4).

(A.1) $\rho(A, A) = 0$,  
(A.2) $\rho(A, \emptyset) = \infty$,  
(A.3) $\rho(A, B \cup C) = \min\{\rho(A, B), \rho(A, C)\}$,  
(A.4) $\rho(A, B) \leq \rho(A, C) + \sup_{C \subseteq \mathcal{P}X} \rho(C, B)$.

**Example 1** Sample approach space.

For a point $x \in X$ and a non-empty set $B$, define the **gap functional** $D_\rho(A, B)$ [8], where

$$D_\rho(A, B) = \begin{cases} \inf \{\rho_{\| \cdot \|}(a, b) : a \in A, b \in B\}, & \text{if } A \text{ and } B \text{ are not empty}, \\ \infty, & \text{if } A \text{ or } B \text{ is empty}. \end{cases}$$

Let $\rho_{\| \cdot \|}$ denote $\| \cdot \| : X \times X :\rightarrow [0, \infty]$ denote the norm on $X \times X$ defined by $\rho_{\| \cdot \|}(\vec{x}, \vec{y}) = \| \vec{x} - \vec{y} \|_1 = \sum_{i=1,n} |x_i - y_i|$.  

**Lemma 1.** $D_\rho_{\| \cdot \|} : \mathcal{P}X \times \mathcal{P}X :\rightarrow [0, \infty]$ satisfies (A.1)-(A.4).

**Proof.** (A.1)-(A.2) are immediate from the definition of $D_\rho_{\| \cdot \|}$. For all $A, B, C \in \mathcal{P}X$, $D_\rho_{\| \cdot \|}$ satisfies (A.3), since

$$D_\rho_{\| \cdot \|}(A, B \cup C) = \inf \{\inf \rho_{\| \cdot \|}(A, B), \inf \rho_{\| \cdot \|}(A, C)\}.$$
Near tolerance rough sets

\[ D_{\rho_{\|\|}}(A, B) \leq \inf_{C \subseteq P_X} \rho_{\|\|}(A, C) + \sup_{C \subseteq P_X} \inf_{C \subseteq P_X} \rho_{\|\|}(C, B) \]
\[ \leq \rho_{\|\|}(A, C) + \sup_{C \subseteq P_X} \rho_{\|\|}(C, B) \]

\[ \square \]

**Theorem 1.** \((X, D_{\rho_{\|\|}})\) is an approach space.

**4 Descriptively Near Sets**

Descriptively near sets are disjoint sets that resemble each other. A feature-based gap functional defined for the norm on a set \(X\) was introduced by J.F. Peters in [21]. In this article, a description-based gap functional \(D_{\Phi_X,\rho_{\|\|}}\) is defined in terms of the Hausdorff lower distance [1] of the norm on \(P\Phi_X \times P\Phi_Y\) for sets \(X,Y \in PX\), i.e. where \(\Phi_X = \{\Phi_1(x), \ldots, \Phi_{|X|}(x)\}\)

\[ D_{\Phi_X,\rho_{\|\|}}(A, B) = \begin{cases} \inf \{\rho(\Phi_X, \Phi_Y)\} & \text{if } \Phi_X \text{ and } \Phi_Y \text{ are not empty,} \\ \infty & \text{if } \Phi_X \text{ or } \Phi_Y \text{ is empty.} \end{cases} \]

**Theorem 2.** \((X, D_{\Phi_X,\rho_{\|\|}})\) is an approach space.

**Proof.** Immediate from the definition of \(D_{\Phi_X,\rho_{\|\|}}\) and Lemma 1.

Given an approach space \(\langle X, \phi \rangle\), define \(\nu : X \times PX \rightarrow [0, \infty]\) by

\[ \nu_{\rho}(x, A) = \inf_{x \in X} \sup_{A \in A} \rho(x, A). \quad (1) \]

The collection \(A \in \mathcal{P}^2X\) is **near** if, and only if \(\nu_{\rho}(x, A) = 0\) for some \(x \in X\) [12]. The function \(\nu_{\rho}\) is called an **approach merotopy** [26]. In the sequel, rewrite (1), replacing \(x \in X\) with \(B \in \mathcal{B} \in \mathcal{P}^2X\) and \(\rho\) with \(D_{\Phi_X,\rho_{\|\|}}\), then,

\[ \nu_{D_{\Phi_X,\rho_{\|\|}}}(A, B) = \inf_{B \in \mathcal{B}} \sup_{A \in A} D_{\Phi_X,\rho_{\|\|}}(B, A). \quad (2) \]

Then the collections \(A,B \in \mathcal{P}^2X\) are **\Phi-near** if, and only if \(\nu_{D_{\Phi_X,\rho_{\|\|}}}(A, B) = 0\) for some \(A \in \mathcal{A}, B \in \mathcal{B}\).

**Definition 1.** **Near Tolerance Rough Sets**

Assume \(A,C \in \mathcal{P}^2X\) are collections of classes in the lower approximation of tolerance rough sets relative to a cover on \(X\) defined by tolerance relation \(\tau\). Then the collections \(A,C\) are **\Phi-near** if, and only if \(\nu_{D_{\Phi_X,\rho_{\|\|}}}(A, C) = 0\) for some \(A \in \mathcal{A}, C \in \mathcal{C}\). When the meaning of the notation is clear from the context, we simply write \(D_{\rho}\) instead of \(D_{\Phi_X,\rho_{\|\|}}\).
Example 2 Sample approach merotopy
Assume that integer values in $A, C$ denote intensities in a pair of greyscale images. Let $\Phi_n(x) = \{\phi\}$, where $\phi(x)$ = greylevel intensity $x$. The set

$$N_\rho(p, \varepsilon) = \{x \in X : \rho(p, x) < \varepsilon\},$$

with $\varepsilon > 0$, is called an open neighbourhood with centre $p$ and radius $\varepsilon$. Using a neighbourhood approach, we get a tolerance class associated with each image pixel (neighbourhood centre). For example, let $\varepsilon = 2$ with $\rho(x, y) = |\phi(x) - \phi(y)| < \varepsilon$,

$$U = \{0, 1, 2, 3, 8, 9, 5, 6, 7, 11, 12\}$$

$$A = \{0, 1, 2, 3, 8\}, \text{ where}$$

$A_0 = N_\rho(0, 2) = \{0, 1\},$

$A_1 = N_\rho(1, 2) = \{1, 2\}, A_2 = N_\rho(2, 2) = \{2, 3, 1\}, A_3 = N_\rho(3, 2) = \{3, 2\},$

$A_4 = N_\rho(8, 2) = \{8, 7, 9\},$

$$C = \{5, 6, 7, 11\}, \text{ where}$$

$C_1 = N_\rho(5, 2) = \{5, 6\}, C_2 = N_\rho(6, 2) = \{6, 5, 7\}, C_3 = N_\rho(7, 2) = \{7, 6\},$

$C_4 = N_\rho(11, 2) = \{11, 12\}.$

Notice that $A$ is a tolerance rough set, since its lower approximation $\Phi_* A$ is

$$\Phi_* A = \{A_1, A_2, A_3\},$$

and class $A_4 \notin A$. Similarly, $C$ is a tolerance rough set, since its lower approximation $\Phi_* C$ is

$$\Phi_* C = \{C_1, C_2, C_3\},$$

and class $C_4 \notin C$.

(step.1) Start by finding the classes in each lower approximation for sets $A, C$.

Every class is a neighbourhood of a pixel.

(step.2) Extract the classes from the lower approximations $\Phi_* A, \Phi_* C$.

(step.3) Next determine the distances between each of the pairs of classes in $\Phi_* A, \Phi_* C$, find $D_\rho(\Phi_* A, \Phi_* C)$, computed in the following way. For illustration, we show one computation.

$$D_\rho(A_1, C_1) = \inf\{\inf\{\rho(1, 5), \rho(1, 6)\},$$

$$\inf\{\rho(2, 5), \rho(2, 6)\}\},$$

$$= \inf\{\inf\{4, 5\}, \inf\{3, 4\}\},$$

$$= \inf\{4, 3\} = 3.$$

$$D_\rho(A_1, C_1) = 3, D_\rho(A_1, C_2) = 3, D_\rho(A_1, C_3) = 4$$

$$D_\rho(A_2, C_1) = 2, D_\rho(A_2, C_2) = 2, D_\rho(A_2, C_3) = 3$$

$$D_\rho(A_3, C_1) = 2, D_\rho(A_3, C_2) = 2, D_\rho(A_3, C_3) = 3.$$
Near tolerance rough sets

(step.4) Now find the sup of each pair of distances.

\[
\sup(D_\rho(A_1, C)) = \sup\{D_\rho(A_1, C_1), D_\rho(A_1, C_2), D_\rho(A_1, C_3)\},
\]
\[
= \sup\{3, 3, 4\} = 4,
\]
\[
\sup(D_\rho(A_2, C)) = 3,
\]
\[
\sup(D_\rho(A_3, C)) = 3.
\]

(step.5) Then, to complete the derivation of the merotopy value, find

\[
\nu_{D_\rho}(\Phi_*A, \Phi_*C) = \inf\{\sup(D_\rho(A_1, C)), \sup(D_\rho(A_2, C)), \sup(D_\rho(A_3, C))\},
\]
\[
= \inf\{4, 3, 3\} = 3.
\]

That is, \(\Phi_*A\) is not near \(\Phi_*C\).

Figure 1: Sample retrieved images for greylevel, edge-orientation features

5 Application: Content-based Image Retrieval

For simplicity, assume that a region of interest in a digital image is tolerance rough set (this is usually the case(see, e.g., [24]). Define an approach space \((U, \rho)\), where \(U\) is a set of digital images and \(\rho\) is a distance function that measures the nearness of pairs of TRSs that are subsets of images in \(U\). We are interested in measuring the distance between digital images (a digital image is viewed as a set of points). The basic approach in the proposed form of Content-based Image Retrieval(CBIR) is to start with a query image \(X \in U\) and use a metric to determine the degree of nearness of \(X\) to images in a set \(Y \in U\). Let \(X,Y\) denote a pair of sets in a metric space, where the distance between sets is measured. The results of experiments with three distance functions from [3] are reported here, namely,

\(M.1\) \(tNM(X,Y)\) (tolerance nearness measure),
\(M.2\) \(tHD(X,Y)\) (tolerance Hausdorff measure),
\(M.3\) \(tHM(X,Y)\) (tolerance Hamming measure).
A detailed explanation of the each of these distance functions is given in [3] and not repeated here.

**Example 2. Sample tNM-based Approach Space**

Assume \( U \) be a non-empty set of images containing subsets that are TRSs. Assume \( A, B \in \mathcal{P}(U) \). Consider the distance \( \rho_{tNM} : \mathcal{P}(U) \times \mathcal{P}(U) \rightarrow [0, \infty] \)

\[
\rho_{tNM}(A, B) = \begin{cases} 
1 - tNM(A, B), & \text{if } A \text{ and } B \text{ are not empty,} \\
\infty, & \text{if } A \text{ or } B \text{ is empty.}
\end{cases}
\]

Then, by definition, \((U, \rho_{tNM})\) is an approach space. Similar distance functions can be defined on two additional approach spaces, \((U, \rho_{tHD}), (U, \rho_{tHM})\)

![Graphs showing CBIR results for greylevel, edge-orientation features](image)

**Figure 2: CBIR results for greylevel, edge-orientation features**

**Example 3. Sample CBIR Results**

This results reported here result from the selection of a leaf image from 186 images of leaves in the CalTech image archive [27]. A sample leaf query image is given in Fig. 1.1. In this sample CBIR experiment, image description is based on two features, namely, average greylevel and average edge orientation of subimages in each image. Out of all of the compared leaves, the leaves descriptively nearest the leaf in the query image are shown in Figs. 1.2-1.5. It is easy to verify that each of the leaves (in isolation as a region-of-interest) in these images (including the query image) are in fact TRSs after a cover on the whole of each image has been determined by a tolerance relation (see [3], e.g., to see how this is done). The plots in Figs. 2.1-2.3 show the CBIR results using each of the three distance functions \( \rho_{tNM}, \rho_{tHD}, \rho_{tHM} \). The tNM distance function gives more trustworthy results. It is left as an open problem to verify whether any collection of the leaf images is, in fact, near a collection of query images, i.e., for any \( \rho \), whether \( \nu_{\rho}(B, A) := \inf_{B \in B} \sup_{A \in A} \rho(B, A) = 0 \), if \( B \in B \in \mathcal{P}^2(U) \) is a set of query images and \( A \in A \in \mathcal{P}^2(U) \) for one of the approach spaces in Example 2.
References


Received: November, 2010