CSE4403 3.0 Introduction to Soft Computing
Tuesdays, Thursdays 10:00 – 11:20 – LAS 3033
Fall Semester, 2013

More Rough Sets

Rough Sets – Dependency of Attributes

Let C and D be subsets of A. We say that D depends on C in a degree k ($0 \le k \le 1$) denoted by $C \rightarrow_k D$ if

$$k = \gamma\,(C, D) = \sum_{X \in U/D} \frac{|\underline{C}(X)|}{|U|}$$

$$k = \gamma\,(C, D) = \frac{|POS_C(D)|}{|U|}$$

where $|POS_C(D)| = \cup_{x\ member\ U/D}\, C\_\,(x)$, called C-positive region of D.

If $k = 1$ we say that D depends totally on C
If $k < 1$ we say that D depends partially (in a degree k) on C.

Rough Sets

U	Category	Major	Birth_Place	Grade
1	PhD	History	Detroit	A
2	MS	Chemistry	Akron	A
3	MS	History	Detroit	C
4	BS	Math	Detroit	B
5	BS	Chemistry	Akron	C
6	PhD	Computing	Cleveland	A
7	BS	Chemistry	Cleveland	C
8	PhD	Computing	Akron	A

U\{Category} = {{1, 6, 8}, {2, 3}, {4, 5, 7}}
U\{Major} = {{1, 3}, {2, 5, 7}, {4}, {6, 8}}
U\{Birth_Place} = {{2, 5, 8}, {1, 3, 4}, {6, 7}}
U\{Grade} = {{1, 2, 6, 8}, {4}, {3, 5, 7}}

C = {Major, Birth_Place}

U\C = {{1, 3}, {2,5}, {4}.{7}, {8}}

U\D = {{1, 2, 6, 8}, {3, 5, 7}}

D = {Grade}

C({1, 2, 6, 8}) = {6, 8}, $\underline{C}$({4}) =

$\underline{C}$({3, 5, 7}) = {7}

$C \rightarrow_k D = |POS_C(D)|/|U| + 4/8 = 1/2$

Rough Sets – Dispensable and Indispensable Attributes

Let S = {U, A = C $\cup$ D be a decision table

Let c be an attribute in C.

Attribute c is dispensable in S if $POS_C(D) = POS_{(C-\{c\})}(D)$

Otherwise, c is indispensable

A decision tale S is independent if all attributes in C are indispensible.

Rough Sets – Reducts and Core

Let S = {U, A = C $\cup$ D be a decision table

A subset R of C is a reduct of C, if

$POS_R(D) = POS_C(D)$ and

S' = (U, R $\cup$ D) is independent, i.e., all attributes in R are indispensable in S'.

Core of C is the set of attributes shared by all reducts of C.

$CORE(C) = \cap\, RED(C)$

where RED(C) is the set of all reducts of C.

Rough Sets – Reducts and Core

U	Category	Major	Birth_Place	Grade
1	PhD	History	Detroit	A
2	MS	Chemistry	Akron	A
3	MS	History	Detroit	C
4	BS	Math	Detroit	B
5	BS	Chemistry	Akron	C
6	PhD	Computing	Cleveland	A
7	BS	Chemistry	Cleveland	C
8	PhD	Computing	Akron	A

U\{Category} = {{1, 6, 8}, {2, 3}, {4, 5, 7}}
U\{Major} = {{1, 3}, {2, 5, 7}, {4}, {6, 8}}
U\{Birth_Place} = {{2, 5, 8}, {1, 3, 4}, {6, 7}}
U\{Grade} = {{1, 2, 6, 8}, {4}, {3, 5, 7}}

$POS_C(D)=\{1,2,3,4,5,6,7,8\}$
$POS_{C1}(D)=\{1,2,3,4,5,6,7,8\}$
$POS_{C2}(D)=\{1,2,3,4,5,6,7,8\}$
$POS_{C3}(D)=\{4,6,7,8\}$

C1 and C2 are reducts of C
C3 is not a reduct of C.

The Core of C is {category}

C = {Category, Major, Birth_Place} D = {Grade}

C_1 = {Category, Major} C_2 = {Category, Birth_Place}

C_3 = {Major, Birth_Place}

U\ C_1 = {{1}, {2}, {3}, {4}, {5, 7}, {6,8}}

U\ C_2 = {{1}, {2}, {3}, {4}, {5}, {6}, (7}, {6,8}}

Rough Sets – Rough Membership Function

Let $S = (U, A)$, $B \subseteq A$ and $X \subseteq U$

Then the rough membership function μB_X for X is a mapping from U to [0, 1]

$$\mu B_X: U \rightarrow [0, 1]$$

For all e in U, the degree of e belongs to X in light of the set of attributes B is defined as

$$\mu B_X (e) = |B(e) \cap X|/B(e)$$

where B(e) denotes the block containing e.

Rough Sets – Properties of Rough Membership Function

P1: $\mu B_X(e) = 1$ iff e in $B_*(X)$

P2: $\mu B_X(e) = 0$ iff e in $U - B^*(X)$

P3: $0 \leq \mu B_X(e) \leq 1$ iff e in $BN_B(X)$

P4: $\mu B_{U-X}(e) = 1 - \mu B_X(e)$ for any e in U

P5: $\mu B_{X \cup Y}(e) \geq \max(\mu B_X(e), \mu B_Y(e))$ for any e in U

P6: $\mu B_{X \cap Y}(e) \leq \min(\mu B_X(e), \mu B_Y(e))$ for any e in U

where $B_*(X)$ is the lower approximation of X in B and $B^*(X)$ is the upper approximation of X in B.

Rough Sets and Fuzzy Sets

Let U be a domain of objects. A fuzzy set X defined on U is characterized by a membership function μ_X:

$$\mu_X : U \rightarrow [0, 1]$$

Let A and B be two fuzzy sets, and

$$\mu_{A \cap B} = \min(\mu_A, \mu_B)$$

$$\mu_{A \cup B} = \max(\mu_A, \mu_B)$$

Let S = (U, R) be an approximate space and X be a subset of U. Define

$\mu X(e) = 1$ if e in _RX

$\mu X(e) = 1/2$ if e in $BN_B(X)$

$\mu X(e) = 1$ if e in -RX where –X is the complement of X

Then the rough membership function cannot be extended to the fuzzy union and intersection of sets.

Rough Sets and Fuzzy Sets

In general:

$$\underline{R}(X \cup Y) \supseteq \underline{R}X \cup \underline{R}Y \quad \text{and} \quad \overline{R}(X \cap Y) \subseteq \overline{R}X \cap \overline{R}Y.$$

The rough membership function will reduce to fuzzy set when

$$\underline{R}(X \cup Y) = \underline{R}X \cup \underline{R}Y \quad \text{and} \quad \overline{R}(X \cap Y) = \overline{R}X \cap \overline{R}Y$$

Rough Sets - Example

Sample Information System

Object	P1	P2	P3	P4	P5
O1	1	2	0	1	1
O2	1	2	0	1	1
O3	2	0	0	1	0
O4	0	0	1	2	1
O5	2	1	0	2	1
O6	0	0	1	2	2
O7	2	0	0	1	0
O8	0	1	2	2	1
O9	2	1	0	2	2
O10	2	0	0	1	0

When the full set of attributes is considered, we see that we have the following seven equivalence classes
{O1, O2}   {O3, O7, O10}
{O4} {O5} {O6} {O8} and {O9}

Thus, the two objects within the first equivalence class, {O1, O2}, cannot be distinguished from one another based on the available attributes, and the three objects within the second equivalence class, {O3, O7,O10}, cannot be distinguished from one another based on the available attributes. The remaining five objects are each discernible from all other objects. The equivalence classes of the P-indiscernibility relation are denoted $[X]_P$.

Rough Sets – Example (cont)

It is apparent that different attribute subset selections will in general lead to different indiscernibility classes. For example, if attribute P=P1 alone is selected, we obtain the following much coarser equivalence class structure:

{O1, O2} {O3, O5, O7, O9, O10} {O4, O6, O8}

Definition of rough set

Let $X \subseteq U$ be a target set that we wish to represent using attribute subset P. That is, we are told that an arbitrary set of objects X comprising a single class, and we wish to express this class (i.e., this subset) using the equivalence classes induced by attribute subset P. In general, X cannot be expressed exactly, because the set may include and exclude objects which are indistinguishable based on attributes .

Rough Sets – Example (cont)

For example, consider the target set X+{O1, O2, O3, O4}, and let attribute subset P={P1, P2, P3, P4, P5}, the full available set of features. It will be noted that the set X cannot be expressed exactly because in $[X]_P$ objects {O3, O7, O10} are indiscernible. Thus, there is no way to represent any set X which includes O3 but excludes objects O7 and O10.

However, the target set can be approximated using only the information contained within P by constructing the P-lower and P-upper approximations of X:

$PX = \{x \mid [x]_P \subseteq X\}$

$PX = \{x \mid [x]_P \cap X \neq \Phi\}$

Rough Sets – Example (cont)

Lower approximation and positive region

The P-lower approximation or positive region is the union of all equivalence classes in $[x]_P$ which are **contained by** (i.e., are subsets of) the target set. In the example, $\underline{P}X = \{O1, O2\} \cup \{O4\}$, the union of the two equivalence classes in $[x]_P$ which are contained in the target set. The lower approximation is the complete set of objects U/P in that can be positively (i.e., unambiguously) classified as belonging to target set .

Rough Sets – Example (cont)

Upper approximation and negative region

The -upper approximation is the union of all equivalence classes in $[x]_P$ which have non-empty intersection with the target set. In the example, $\overline{P}X = \{O1, O2\} \cup \{O4\} \cup \{O3, O7, O10\}$, the union of the three equivalence classes in $[x]_P$ that have non-empty intersection with the target set. The upper approximation is the complete set of objects that in U/P that cannot be positively (i.e., unambiguously) classified as belonging to the complement of the target set X. In other words, the upper approximation is the complete set of objects that are possibly members of the target set X.

The set U - $\overline{P}$X therefore represents the negative region, containing the set of objects that can be definitely ruled out as members of the target set.

Rough Sets – Example (cont)

The boundary region, given by set difference PX = - PX, consists of those objects that can neither be ruled in nor ruled out as members of the target set .

In summary, the lower approximation of a target set is a *conservative* approximation consisting of only those objects which can positively be identified as members of the set. (These objects have no indiscernible "clones" which are excluded by the target set.) The upper approximation is a *liberal* approximation which includes all objects that might be members of target set. (Some objects in the upper approximation may not be members of the target set.) From the perspective of U/P, the lower approximation contains objects that are members of the target set with certainty (probability = 1), while the upper approximation contains objects that are members of the target set with nonzero probability (probability > 0).

Rough Sets – Example (cont)

The tuple $\langle \underline{P}X, \overline{P}X \rangle$ composed of the lower and upper approximation is called a rough set. Thus, a rough set is composed of two crisp sets, one representing a lower boundary of the target set X, and one representing an upper boundary of the target set X.

The accuracy of the rough set representation of the set can be given (Pawlak 1991) by the following:

$$\alpha_p(X) = \underline{P}X / PX$$

That is, the accuracy of the rough set representation of X is the ratio of the number of objects which can be *positively* placed in X to the number of objects that can be *possibly* be placed in X, This provides a measure of how closely the rough set is approximating the target set. Clearly, when the upper and lower approximations are equal (i.e., boundary region empty), then , and the approximation is perfect. Whenever the lower approximation is empty, the accuracy is zero (regardless of the size of the upper approximation).

Rough Sets – Example (cont)

An interesting question is whether there are attributes in the information system (attribute-value table) which are more important to the knowledge represented in the equivalence class structure than other attributes. Often we wonder whether there is a subset of attributes which by itself can fully characterize the knowledge in the database. Such an attribute set is called a reduct.

Formally (Ziarko & Shan 1995), a reduct is a subset of attributes $RED \subseteq P$ such that

$[x]_{RED} = [x]_p$, that is, the equivalence classes induced by reduced attribute set RED is the same as the equivalence class structure induced by full attribute set P.

Attribute set RED is minimal in the sense that $x_{(RED-a)}$ $[x]_p$ for any attribute $A \in RED$. In other words, no attribute can be removed from set RED without changing the equivalence classes $[x]_p$.

Rough Sets – Example (cont)

A reduct can be thought of as a sufficient set of features; sufficient, that is, to represent the category structure. In the example table above, attribute set {P3, P4, P5} is a reduct. The information system projected on just these attributes possesses the same equivalence class structure as that expressed by the full attribute set:

{O1, O2}   {O3, O7, O10}   {O4}   {O5}   {O6}   {O8}   {O9}

Attribute set {P3, P4, P5} is a legitimate reduct because eliminating any of these attributes causes a collapse of the equivalence class structure, with the result that $x_{RED} \neq [x]_p$

The reduct of an information system is **not unique**. There may be many subsets of attributes which preserve the equivalence class structure (i.e., the knowledge) expressed in the information system. In the example information system above, another reduct is {P1, P2, P5}, producing the same equivalence class structure as $[x]_p$

Rough Sets – Example (cont)

The set of attributes which is common to all reducts is called the core. The core is the set of attributes which is possessed by every legitimate reduct, and therefore consists of attributes which cannot be removed from the information system without causing collapse of the equivalence class structure. The core may be thought of as the set of necessary attributes; necessary, that is, for the category structure to be represented. In the example, the only such attribute is {P5}. Any one of the other attributes can be removed in isolation without damaging the equivalence class structure, and hence these are all dispensable. However, removing {P5} in isolation *does* change the equivalence class structure, and thus {P5} is the indispensable of this information system, and hence the core.

Rough Sets – Example (cont)

It is possible for the core to be empty, which means that there is no indispensable attribute. Any single attribute in the information system can be deleted without altering the equivalence class structure. In such cases, there is no *essential* or necessary attribute which is required for the class structure to be represented.

Concluding Remarks

Put up in a place
where it's easy to see
the cryptic admonishment T. T. T.

When you feel how depressingly slowly you climb,
it's well to remember that Things Take Time.