On Optimal Approximations for Rough Sets

Ryszard Janicki

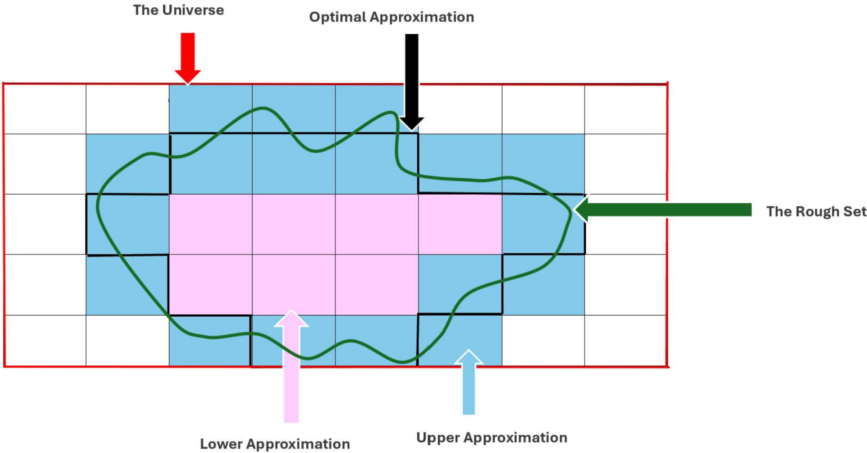
Department of Computing and Software, McMaster University, Hamilton, Ontario, Canada

IJCRS'2024
Halifax, Nova Scotia, Canada
May 17-20, 2024

- Classical Rough Sets consider only lower and upper approximations, however the concept of an approximation is not restricted only to lower and upper approximations.
- Consider the well known *linear least squares approximation* of points in the two dimensional plane (credited to C. F. Gauss, 1795). Here we know or assume that the points should be on a straight line and we are trying to find the line that fits the data best.
- However, this is not the case of an upper, or lower approximation in the sense of Rough Sets.
- The cases like the linear least squares approximation assume that there is a well defined concept of *similarity* (or *distance*) and some techniques for finding *maximal similarity* (*minimal distance*) between entities and their approximations.

- In 2013 and 2016 the concept of **optimal approximation** has been added to *standard Rough sets* and discussed in some detail.
- In 2017 the concept of **optimal approximation** was extended to the *Rough sets defined by coverings*.

Optimal Approximation - Intuition



- Optimal Approx = Lower Approx $\cup$ some elements from (Upper Approx $-$ Lower Approx)
- The problem is 'which ones?'

- The word ‘optimal’ suggests some numerical calculations and comparisons.
- We need some proper definition of *similarity measure* (or index) for sets, including some axioms.
- Any formal definition of similarity must involve the concept of *measure*, either abstract or specific.
- To tackle the problem that different similarity indexes may produce identical results we introduced the concept of *consistent similarities*.
- In some contexts, similarity can be seen as an inverse of *distance*.

Measure

- ' Measure' is a abstract generalization of such concepts as cardinality, length, area, volume, etc.
- Let U be a set, called *universe*, and let $\mathbb{R}$ be the set of real numbers.
- A function $\mu : 2^U \to \mathbb{R}$ is a *finite measure* over 2^U if it satisfies:
  1. for all $X \subseteq U$, $0 \leq \mu(X) < \infty$,
  2. $\mu(\emptyset) = 0$,
  3. if $X_i \subseteq U$ for $i = 1, \ldots, \infty$ and $X_i \cap X_j = \emptyset$ if $i \neq j$, then

$$\mu\left(\bigcup_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} \mu(X_i).$$

- A measure μ is *null set free* if $\mu(X) = 0 \iff X = \emptyset$.
- One of the most popular measures for finite sets is *cardinality*, i.e. $\mu(X) = |X|$.

- Let X be a set, and let $\mathbb{R}$ be the set of real numbers.
- A function $dist : X \to \mathbb{R}$ is a *distance* if and only if
  1. $dist(A,B) \geq 0$,
  2. $dist(A,B) = 0 \iff A = B$,
  3. $dist(A,B) = dist(B,A)$,
  4. $dist(A,C) \leq dist(A,B) + dist(B,C)$, i.e *triangle inequality*.

  for all $A, B, C \in X$.

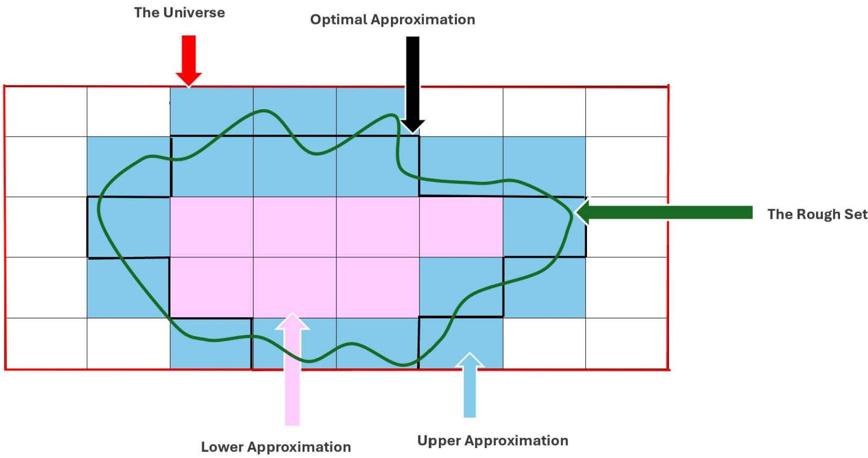
Similarity Axioms

- As opposed to the orthogonal concept of a *distance*, the concept of *similarity* does not have standard indisputable axiomatization.
- Depending on the area of application, some desirable properties may vary, sometimes substantially.
- In our approach we assume that the *similarity* is any (total) function $sim : 2^U \times 2^U \to [0,1]$ satisfying the following five axioms:

$$
\begin{array}{ll}
S1 \text{ (Maximum)}: & sim(A,B) = 1 \iff A = B, \\
S2 \text{ (Symmetry)}: & sim(A,B) = sim(B,A), \\
S3 \text{ (Minimum)}: & sim(A,B) = 0 \iff A \cap B = \emptyset, \\
S4 \text{ (Inclusion)}: & \text{if } a \in B \setminus A \text{ then } sim(A,B) < sim(A \cup \{a\}, B), \\
S5 \text{ (Exclusion)}: & \text{if } a \notin A \cup B \text{ and } A \cap B \neq \emptyset \text{ then} \\
& \quad sim(A,B) > sim(A \cup \{a\}, B) \\
S5' \text{ (Weak Exclusion)}: & \text{if } a \notin A \cup B \text{ then } sim(A,B) \geq sim(A \cup \{a\}, B)
\end{array}
$$

- A similarity *sim* is *metrical* if the function $dist(A,B) = 1 - sim(A,B)$ is a proper *distance*.

- U is a finite and non-empty *universe* of elements, $E \subseteq U \times U$ be an *equivalence relation*.
- The pair $AS = (U, E)$ is called (Pawlak) approximation space.
- Let U/E denote the set of all equivalence classes of E. Note that U/E is also a *partition* of U.
- The elements of $\mathrm{ESets} = U/E$ are called *elementary sets*.
- A set $A \subseteq U$ is *definable* if it is a union of some equivalence classes of the equivalence relation E. Let DSets denote the family of all definable sets.
- Lower Approximation: $\underline{\mathbf{A}}(X) = \bigcup\{\mathtt{A} \mid \mathtt{A} \in \mathrm{Esets} \wedge \mathtt{A} \subseteq X\}$.
- Upper Approximation: $\overline{\mathbf{A}}(X) = \bigcup\{\mathtt{A} \mid \mathtt{A} \in \mathrm{Esets} \wedge \mathtt{A} \cap X \neq \emptyset\}$.
- For every $\mathrm{A} \in \mathrm{DSets}$, the set $\mathrm{esets}(\mathrm{A}) = \{\mathtt{B} \mid \mathtt{B} \subseteq \mathrm{A} \wedge \mathtt{B} \in \mathrm{ESets}\}$, is *its collection of elementary sets*.
- Border: $\mathfrak{B}(X) = \{\mathtt{B} \mid \mathtt{B} \in \mathrm{esets}(\overline{\mathbf{A}}(X)) \setminus \mathrm{esets}(\underline{\mathbf{A}}(X))\} \subseteq \mathrm{ESets}$.
- Border Sets: $\mathbb{B}(X) = \{\mathrm{A} \mid \mathrm{A} \subseteq \overline{\mathbf{A}}(X) \setminus \underline{\mathbf{A}}(X) \wedge \mathrm{A} \in \mathrm{DSets}\} \subseteq \mathrm{DSets}$.



- Boarder: All blue rectangles.
- Boarder Sets: Unions of blue rectangles.
- Optimal Approx = Lower Approx $\cup$ some Boarder Set

Optimal Approximation

Definition

For every set $X \subseteq U$*, a definable set* $\mathrm{O} \in \mathrm{DSets}$ *is an* optimal approximation *of* X *(w.r.t. a given similarity measure sim) if and only if:*

$$sim(X, \mathrm{O}) = \max_{\mathrm{A} \in \mathrm{DSets}} (sim(X, \mathrm{A}))$$

The set *of all optimal approximations of* X *is denoted by* $\mathrm{Opt}_{sim}(X)$. □

Proposition

For every set $X \subseteq U$*, and every* $\mathrm{O} \in \mathrm{Opt}_{sim}(X)$*:*

$$\underline{\mathbf{A}}(X) \subseteq \mathrm{O} \subseteq \overline{\mathbf{A}}(X)$$ □

Popular (Symmetric) Similarity Measures

- *Marczewski-Steinhaus index* (1958): $sim_{MS}(X,Y) = \dfrac{\mu(X \cap Y)}{\mu(X \cup Y)}$.
  First proposed by Jaccard in 1901 with $\mu(X) = |X|$.
- *Dice-Sørensen index* (1945, 1957): $sim_{DS}(X,Y) = \dfrac{2\mu(X \cap Y)}{\mu(X) + \mu(Y)}$.
  Initially proposed with $\mu(X) = |X|$.
- *Braun-Blanquet index* (1928) : $sim_{BB}(X,Y) = \dfrac{\mu(X \cap Y)}{\max(\mu(X), \mu(Y))}$.
  Initially proposed with $\mu(X) = |X|$.
- *Symmetric Tversky index* (1977), it has a parameter α:
  $sim_T^{\alpha}(X,Y) = \dfrac{\mu(X \cap Y)}{\mu(X \cap Y) + \alpha\mu(X \setminus Y) + \alpha\mu(Y \setminus X)}$.
  Initially proposed with $\mu(X) = |X|$.
  Note that for $\alpha = 1$, $sim_T^{\alpha}(X,Y) = sim_{MS}(X,Y)$
  while for $\alpha = 0.5$, $sim_T^{\alpha}(X,Y) = sim_{DS}(X,Y)$.
- Only Marczewski-Steinhaus index is *metrical*!

Asymmetric Similarity Measures

- *Tversky index* (1977), it has a parameters α, β:
  $sim_T^{\alpha,\beta}(X,Y) = \dfrac{\mu(X \cap Y)}{\mu(X \cap Y) + \alpha\mu(X \setminus Y) + \beta\mu(Y \setminus X)}$.
- Tversky index is an asymmetric by design similarity index on sets that compares a variant to a prototype.
- If we consider X to be the prototype and Y to be the variant, then α corresponds to the weight of the prototype and β corresponds to the weight of the variant. For the interpretation of X and Y as prototype and variant, α usually differs from β.
- However for the interpretations used in our approach, the case $\alpha \neq \beta$ is debatable, however it might be worth some consideration.

Consistent Similarity Measures

Definition

We say that two similarity indexes sim_1 and sim_2 satisfying axioms S1–S5 are consistent *if for all sets $A, B, C \subseteq U$,*

$$sim_1(A,B) < sim_1(A,C) \iff sim_2(A,B) < sim_2(A,C). \qquad \square$$

Theorem

If sim_1 and sim_2 are consistent then for each $X \subseteq U$,

$$\mathrm{Opt}_{sim_1}(X) = \mathrm{Opt}_{sim_2}(X). \qquad \square$$

Theorem

1. *The following three similarity indexes are consistent: Marczewski-Steinhaus, Symmetric Tversky, and Dice-Sørensen.*
2. *The Marczewski-Steinhaus index and Braun-Blanquet index are not consistent.* $\square$

Finding Optimal Approximations

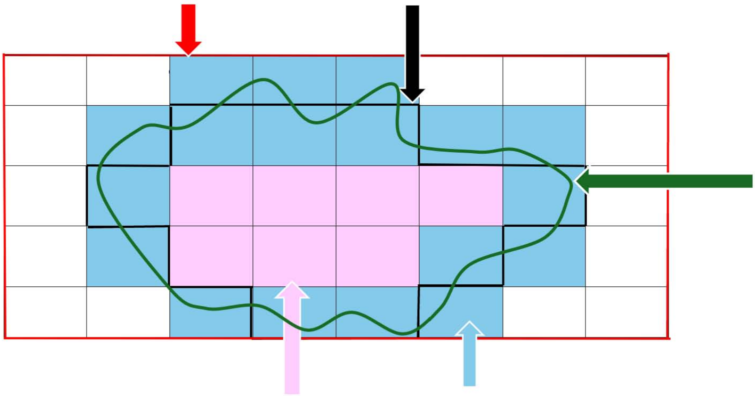
- Let *sim* be a given similarity measure.
- We know that some optimal approximation O does exist, and that $\underline{\mathbf{A}}(X) \subseteq \mathrm{O} \subseteq \overline{\mathbf{A}}(X)$.
- We also know that any optimal approximation of X, is the union of the lower approximation of X and some element $\mathrm{A} \in \mathbb{B}(X) \cup \{\emptyset\}$.

Definition

Let $X \subseteq U$, and $\mathrm{O} \in \mathrm{DSets}$. We say that O is an ***intermediate approximation*** of X, if

$$\underline{\mathbf{A}}(X) \subseteq \mathrm{O} \subseteq \overline{\mathbf{A}}(X)$$

The set of all intermediate approximations of X will be denoted by $\mathbf{IA}(X)$. □



- An *intermediate approximation* is all ‘pink rectangles’ ∪ any collection of ‘blue rectangles’.

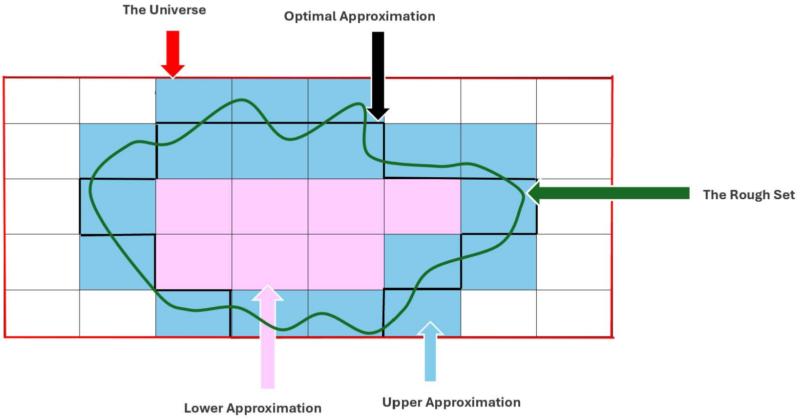
Ratio Common/Distinct

Definition (Ratio common/distinct)

For every $X, Y \subseteq U$, such that $X \setminus Y \neq \emptyset$, we define the index $\rho(X, Y)$, called the *ratio of common to distinct elements*, as follows

$$\rho(X, Y) = \frac{\mu(X \cap Y)}{\mu(X \setminus Y)}.$$

Note that $\rho(X, Y)$ is sound only if μ is finite and null-free. $\square$



- Assume $\mu(X) =$ area of X.
- Y = The Rough Set
- X = Blue rectangle pointed by green arrow: $\rho(X, Y) > 1$.
- X = Blue rectangle pointed by black arrow: $\rho(X, Y) > 1$.
- X = Blue rectangle pointed by blue arrow: $\rho(X, Y) < 1$.
- X = Blue rectangle pointed by red arrow: $\rho(X, Y) < 1$.

Theoretical Foundations

- Now, we assume that our similarity measure is Marczewski-Steinhaus index $sim_{MS}(X,Y) = \frac{\mu(X \cap O)}{\mu(X \cup O)}$.
- It is fairly general, often used, has convincing intuition, and it is consistent with two other popular indexes.

Lemma

Let $X \subseteq U$, $O \in \mathbf{IA}(X)$, $A, B \in \mathbb{B}(X)$, $A \cap O = \emptyset$, *and* $B \subseteq O$. *Then*

1. $sim_{MS}(X, O \cup A) \geq sim_{MS}(X, O) \iff \rho(A, X) \geq sim_{MS}(X, O)$
2. $sim_{MS}(X, O \setminus B) \leq sim_{MS}(X, O) \iff \rho(B, X) \leq sim_{MS}(X, O)$ $\square$

- if $O \in \mathbf{Opt}(X)$, then $O = \underline{\mathbf{A}}(X) \cup B_1 \cup \ldots \cup B_k$, for some k, where each $B_i \in \mathfrak{B}(X)$, $i = 1, \ldots, k$. The above lemma allows us to explicitly define these $B_i \in \mathfrak{B}(X)$ components.

Lemma

Let $X \subseteq U$, $\mathrm{O} \in \mathbf{IA}(X)$, $\mathrm{A}, \mathrm{B} \in \mathbb{B}(X)$, $\mathrm{A} \cap \mathrm{O} = \emptyset$, *and* $\mathrm{B} \subseteq \mathrm{O}$. *Then*

1. $sim_{MS}(X, \mathrm{O} \cup \mathrm{A}) \geq sim_{MS}(X, \mathrm{O}) \iff \rho(\mathrm{A}, X) \geq sim_{MS}(X, \mathrm{O})$
2. $sim_{MS}(X, \mathrm{O} \setminus \mathrm{B}) \leq sim_{MS}(X, \mathrm{O}) \iff \rho(\mathrm{B}, X) \leq sim_{MS}(X, \mathrm{O})$ $\square$

- We know from if $\mathrm{O} \in \mathbf{Opt}(X)$, then $\mathrm{O} = \underline{\mathbf{A}}(X) \cup \mathtt{B}_1 \cup \ldots \cup \mathtt{B}_k$, for some k, where each $\mathtt{B}_i \in \mathfrak{B}(X)$, $i = 1, \ldots, k$.
- The above lemma allows us to explicitly define these $\mathtt{B}_i \in \mathfrak{B}(X)$ components.

Theorem

For every $X \subseteq U$, *the following two statements are equivalent:*

1. $\mathrm{O} \in \mathbf{Opt}(X)$
2. $\mathrm{O} \in \mathbf{IA}(X) \wedge$
   $\left(\forall \mathtt{B} \in \mathfrak{B}(X).\ \mathtt{B} \subseteq \mathrm{O} \iff \rho(\mathtt{B}, X) = \frac{\mu(\mathtt{B} \cap X)}{\mu(\mathtt{B} \setminus X)} \geq sim_{MS}(X, \mathrm{O})\right)$. $\square$

Towards a Greedy Algorithm

- Assume that $r = |\mathfrak{B}(X)|$, $\mathfrak{B}(X) = \{\mathtt{B}_1, \ldots, \mathtt{B}_r\}$ and also
$$i \leq j \iff \rho(\mathtt{B}_i, X) \geq \rho(\mathtt{B}_j, X).$$

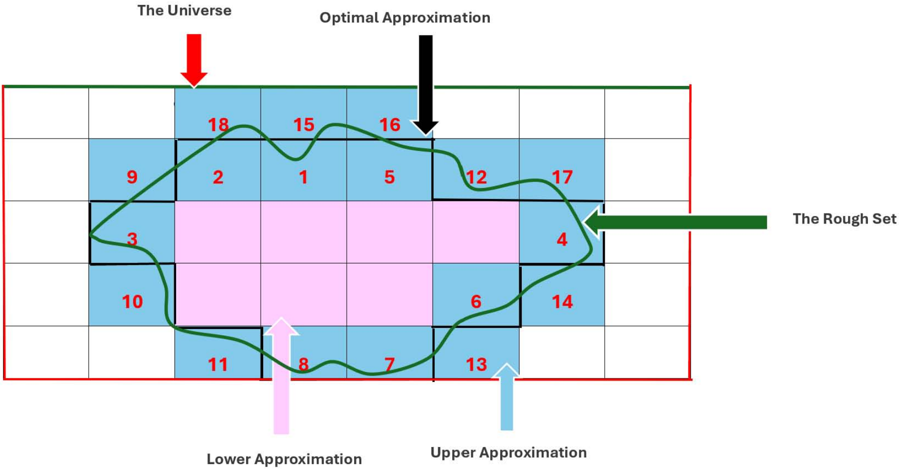
Definition

Let $\mathrm{O}_0, \mathrm{O}_1, \ldots, \mathrm{O}_r \in \mathbf{IA}(X)$ be the sequence of intermediate approximations of X defined for $i = 0, \ldots, r-1$ as follows: $\mathrm{O}_0 = \underline{\mathbf{A}}(X)$ and

$$\mathrm{O}_{i+1} = \begin{cases} \mathrm{O}_i \cup \mathtt{B}_{i+1} & \text{if } sim_{MS}(X, \mathrm{O}_i \cup \mathtt{B}_{i+1}) \geq sim_{MS}(X, \mathrm{O}_i) \\ \mathrm{O}_i & \text{otherwise.} \end{cases}$$

Clearly $\underline{\mathbf{A}}(X) = \mathrm{O}_0 \subseteq \mathrm{O}_1 \subseteq \ldots \subseteq \mathrm{O}_r \subseteq \overline{\mathbf{A}}(X)$. □

- We claim that at least one of these O_i's is an optimal approximation.



- Let X be our green rough set,
- $O_8 = \underline{\mathbf{A}}(X) \cup \mathrm{B}_1 \cup \ldots \cup \mathrm{B}_8$ is an optimal approximation, the only one in this case.

Main Theorem

Theorem

For every $X \subseteq U$, we set $r = |\mathfrak{B}(X)|$, and we have

1. $sim_{MS}(X, \mathrm{O}_{i+1}) \geq sim_{MS}(X, \mathrm{O}_i)$, *for* $i = 0, \ldots, r-1$.
2. *If* $\rho(\mathbb{B}_1, X) \leq sim_{MS}(X, \underline{\mathbf{A}}(X))$ *then* $\underline{\mathbf{A}}(X) \in \mathbf{Opt}(X)$.
3. *If* $\rho(\mathbb{B}_r, X) \geq sim_{MS}(X, \overline{\mathbf{A}}(X))$ *then* $\overline{\mathbf{A}}(X) \in \mathbf{Opt}(X)$.
4. *If* $sim_{MS}(X, \mathrm{O}_p) \leq \rho(\mathbb{B}_p, X)$ *and* $sim_{MS}(X, \mathrm{O}_{p+1}) > \rho(\mathbb{B}_{p+1}, X)$, *then* $\mathrm{O}_p \in \mathbf{Opt}(X)$, *for* $p = 1, \ldots, r-1$.
5. *If* $\mathrm{O}_p \in \mathbf{Opt}(X)$, *then* $\mathrm{O}_i = \mathrm{O}_p$ *for all* $i = p+1, \ldots, r$. *In particular* $\mathrm{O}_r \in \mathbf{Opt}(X)$.
6. $\mathrm{O} \in \mathbf{Opt}(X) \implies \mathrm{O} \subseteq \mathrm{O}_p$, *where p is the smallest one from (5).*

Algorithm (Finding the Greatest Optimal Approximation)

Let $X \subseteq U$.

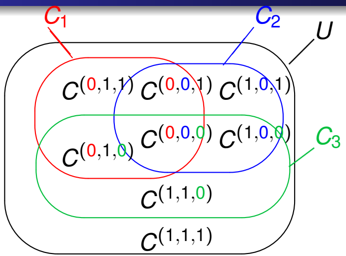
1. *Construct* $\underline{\mathbf{A}}(X)$, $\overline{\mathbf{A}}(X)$, *and* $\mathfrak{B}(X)$. *Assume* $r = |\mathfrak{B}(X)|$.
2. *For each* $\mathtt{B} \in \mathfrak{B}(X)$, *calculate* $\rho(\mathtt{B}, X) = \frac{\mu(\mathtt{B} \cap X)}{\mu(\mathtt{B} \setminus X)}$.
3. *Order* $\rho(\mathtt{B}, X)$ *in decreasing order and number the elements of* $\mathfrak{B}(X)$ *by this order, so* $\mathfrak{B}(X) = \{\mathtt{B}_1, \ldots, \mathtt{B}_r\}$ *and* $i \leq j \iff \rho(\mathtt{B}_i, X) \geq \rho(\mathtt{B}_j, X)$.
4. *If* $\rho(\mathtt{B}_1, X) \leq sim_{MS}(X, \underline{\mathbf{A}}(X))$ *then* $\mathrm{O} = \underline{\mathbf{A}}(X)$.
5. *If* $\rho(\mathtt{B}_r, X) \geq sim_{MS}(X, \overline{\mathbf{A}}(X))$ *then* $\mathrm{O} = \overline{\mathbf{A}}(X)$.
6. *Calculate* O_i *from* $i = 0$ *until* $sim_{MS}(X, \mathrm{O}_{p+1}) > \rho(\mathtt{B}_{p+1}, X)$, *for* $p = 0, \ldots, r-1$, *and set* $\mathrm{O} = \mathrm{O}_p$.

From the main theorem we have that O *is the greatest optimal approximation, i.e.* $\mathrm{O} \in \mathbf{Opt}(X)$, *and for all* $\mathrm{O}' \in \mathbf{Opt}(X)$, $\mathrm{O}' \subseteq \mathrm{O}$. *We also know that* $sim_{MS}(X, \mathrm{O}') = sim_{MS}(X, \mathrm{O})$. □

- This *greedy* algorithm has a complexity of $C_1 + C_2 + \mathrm{O}(r\log r)$, where C_1 is the complexity of constructing $\underline{\mathbf{A}}(X)$, $\overline{\mathbf{A}}(X)$, and $\mathfrak{B}(X)$; while C_2 is the complexity to assign $\mu(x)$ for each $x \in U$.
- We know that $C_1 = \mathrm{O}(|U|^2)$, and clearly $C_2 = \mathrm{O}(|U|)$.
- The most crucial line of the algorithm, line (6), runs in $\mathrm{O}(r)$, but line (3) involves sorting which has complexity $\mathrm{O}(rlogr)$. Since $r < |U|$, the total complexity is $\mathrm{O}(|U|^2)$.
- The algorithm gives us the *greatest optimal approximation* O, however the whole set $\mathbf{Opt}(X)$ can easily be derived from O just by subtracting appropriate elements of $\mathfrak{B}(X)$.

- The algorithm can also be used for *any* similarity measure *sim* that is *consistent* with the Marczewski-Steinhaus index sim_{MS}.
- The algorithm is based on the two fundamental assumptions: the set ESets is a partition, and the similarity measure is sim_{MS}, so the ration ρ can be used to pick appropriate elements from the border $\mathfrak{B}$.
- If any of the is not satisfied, the algorithm cannot be used.
- So, what about Rough Sets induced by Coverings?
- It turns out our approach can also be applied to some Rough Sets induced by Coverings.
- The idea is to replace coverings by appropriate equivalent partitions.

Components



- All eight components for a given set U and subsets $\mathbb{C} = \{C_1, C_2, C_3\}$.
- In general: $\mathbb{C} = \{C_i \mid i = 1 \ldots n\}$ and $C^{(i_1,\ldots,i_n)} = C_1^{i_1} \cap \ldots \cap C_n^{i_n}$, where $i_k = 0, 1$, and $C_i^0 = C_i$, $C_i^1 = U \setminus C_i$ for $k = 1, 2, \ldots, n$.
- Let: $\text{comp}(\mathbb{C}) = \{C^{(i_1,\ldots,i_n)} \mid i_k = 1, 2,\ k = 1, 2, \ldots, n,\ C^{(i_1,\ldots,i_n)} \neq \emptyset\}$.
- $\text{cov}(\mathbb{C}) = \mathbb{C} \cup \{U \setminus \bigcup_{i=1}^n C_i\}$ is a covering of U.
- We might assume $U = \bigcup_{i=1}^n C_i$, if convenient.
- $\text{comp}(\mathbb{C})$ is partition.
- We can derive $\text{comp}(\mathbb{C})$ from $\text{cov}(\mathbb{C})$ and vice versa, so we may consider them as equivalent.

- Let U be a set and let $\mathbb{C}$ be a nonempty family of nonempty subsets of U.
- A non-empty set $X \subseteq U$ is *definable* by $\mathbb{C}$ if it can be constructed from the elements of $\mathbb{C}$ by means of set operations $\cup$, $\cap$ and $\setminus$.
- The family of all sets *definable* by $\mathbb{C}$ will be denoted by definable($\mathbb{C}$).
- $\mathbb{C}$ *does not have to* be a covering.

Theorem

Let U be a set and $\mathbb{C}$ a family of nonempty subsets of U.

1. $$U = \bigcup_{C^{(\alpha)} \in \text{comp}(\mathbb{C})} C^{(\alpha)}.$$
2. *For all $C^{(\alpha)}, C^{(\beta)} \in \text{comp}(\mathbb{C})$, we have* $C^{(\alpha)} \cap C^{(\beta)} = \emptyset \iff \alpha \neq \beta$.
3. *Every set $X \in \text{definable}(\mathbb{C})$ is a union of some components of $\mathbb{C}$, i.e.*
$$X \in \text{definable}(\mathbb{C}) \iff \exists C^{(\alpha_1)}, \ldots, C^{(\alpha_k)} \in \text{comp}(\mathbb{C}).\ X = C^{(\alpha_1)} \cup \ldots \cup C^{(\alpha_k)}.$$

□

Rough Sets Defined By Coverings

- There are several, similar but not identical models.
- U is a finite and non-empty universe of elements, and $\mathbb{C}$ is its *covering*.
- The pair $\mathrm{AS} = (U, \mathbb{C})$ is an *covering approximation space*.
- For every $x \in U$, the set $N(x) = \bigcap\{C \in \mathbb{C} \mid x \in C\}$ is called the *neighborhood* of an element $x \in U$.
- *Elementary sets*: $\mathrm{ESets} = \{N(x) \mid x \in U\}$.
- *Definable sets*: DSets, are set unions of elementary sets.
- *Lower approximations* are defined differently, for example:
  (a) $\underline{\mathbf{A}}_a(X) = \bigcup\{C \in \mathbb{C} \mid C \subseteq X\}$.
  (b) $\underline{\mathbf{A}}_b(X) = \{x \mid N(x) \subseteq X\}$.
- *Upper approximation* are defined differently, for example:
  (a) $\mathbf{A}^+(X)\overline{\mathbf{A}}_a(X) = \underline{\mathbf{A}}_a(X) \cup \bigcup\{N(x) \mid x \in X \setminus \underline{\mathbf{A}}_a(X)\}$.
  (b) $\overline{\mathbf{A}}_b(X) = \{x \mid N(x) \cap X \neq \emptyset\}$.
- *Optimal approximation* of X, w.r.t similarity measure:

$$sim(X, \mathrm{O}) = \max_{\mathrm{A} \in \mathrm{DSets}} (sim(X, \mathrm{A})).$$

An efficient algorithm from classical rough sets does not work.

When Partitions Are Derived From Coverings

- *U* is a finite and non-empty universe of elements, and $\mathbb{C}$ is its *covering*.
- The pair $AS = (U, \mathbb{C})$ is an *covering approximation space*.
- *comp*$(\mathbb{C})$ is the set of all components of $\mathbb{C}$.
- *Elementary sets*: $\text{ESets} = \mathbb{C}$.
- *Definable sets*: $\text{DSets} = \text{definable}(\mathbb{C})$, i.e. all sets that can be derived from $\mathbb{C}$ by applying operations $\cup$, $\cap$ and $\setminus$.
- *Lower approximation* of X:
$$\underline{\mathbf{A}}^c(X) = \bigcup\{C^{(\alpha)} \mid C^{(\alpha)} \in \text{comp}(\mathbb{C}) \wedge C^{(\alpha)} \subseteq X\}.$$
- *Upper approximation* of X:
$$\overline{\mathbf{A}}^c(X) = \bigcup\{C^{(\alpha)} \mid C^{(\alpha)} \in \text{comp}(\mathbb{C}) \wedge C^{(\alpha)} \cap X \neq \emptyset\}.$$
- *Optimal approximation* of X, w.r.t similarity measure:
$$sim(X, \mathrm{O}) = \max_{\mathrm{A} \in \text{DSets}} (sim(X, \mathrm{A})).$$

An efficient algorithm from classical rough sets DOES work.

- $\underline{\mathbf{A}}^c(X), \overline{\mathbf{A}}^c(X)$ are **never less tight** than appropriate approximations derived directly from coverings.

Final Comment

- One intriguing problem, still open, is the relationship between the measure μ and the equivalence relation E.
- The relation E represents the knowledge of an approximation space (U, E) and defines the set DSets.
- Our definition of optimal approximation indicates that an optimal approximation is a function of both the measure μ and the equivalence relation E, but how μ and E relate is an open problem.

THANK YOU

QUESTIONS?