Introduction

- Main goal of rough sets - induction of approximations of concepts.

- It constitutes sound basis for knowledge discovery in databases
  - offers mathematical tools to discover patterns hidden in data.

- It can be used for feature selection, feature extraction, data reduction, decision rule generation, and pattern extraction (templates, association rules), clustering, etc.

- Identifies partial or total dependencies in data, eliminates redundant data, gives approach to null values, missing data, dynamic data and others.

Basic Concepts of Rough Sets

- Information/Decision Systems (Tables)
- Indiscernibility
- Set Approximation
- Reducts and Core
- Rough Membership
- Dependency of Attributes

Information Systems/Tables

U	A1	A2
x1	16-30	50
x2	16-30	0
x3	31-45	1-25
x4	31-45	1-25
x5	46-60	26-49
x6	16-30	26-49
x7	46-60	26-49

- *IS* is a pair (U, A)
- U is a non-empty finite set of objects.
- A is a non-empty finite set of attributes such that $a : U \rightarrow V_a$ for every $a \in A$.
- V_a is called the value set of a.

Decision Systems/Tables: Issues

U	A1	A2	Walk
x1	16-30	50	yes
x2	16-30	0	no
x3	31-45	1-25	no
x4	31-45	1-25	yes
x5	46-60	26-49	no
x6	16-30	26-49	yes
x7	46-60	26-49	no

- *DS*: $T = (U, A \cup \{d\})$
- Elements of A={A1, A2} are called the *condition attributes*.
- $d \notin A$ is the *decision attribute* (instead of one we can consider more decision attributes).
- The same or indiscernible objects may be represented several times.
- Some of the attributes may be superfluous.

An Example of Indiscernibility

U	A1	A2	Walk
x1	16-30	50	yes
x2	16-30	0	no
x3	31-45	1-25	no
x4	31-45	1-25	yes
x5	46-60	26-49	no
x6	16-30	26-49	yes
x7	46-60	26-49	no

- Non-empty subsets of condition attributes are {A1}, {A2}, and {A1, A2}.
- *IND*({A1}) = {{x1,x2,x6}, {x3,x4}, {x5,x7}}
- *IND*({A2}) = {{x1}, {x2}, {x3,x4}, {x5,x6,x7}}
- *IND*({A1, A2}) = {{x1}, {x2}, {x3,x4}, {x5,x7}, {x6}}.

Indiscernibility

- Equivalence relation – a binary relation $R \subseteq X \times X$ which is
  - reflexive (xRx for any object x),
  - symmetric (if xRy then yRx), and
  - transitive (if xRy and yRz then xRz).

- The *equivalence class* $[x]_R$ of an element $x \in X$ consists of all objects $y \in X$ such that xRy.

- Let $IS = (U, A)$ be an *information system*, then with any $B \subseteq A$ there is an associated equivalence relation:

$$IND_{IS}(B) = \{(x, x') \in U^2 \mid \forall a \in B, a(x) = a(x')\}$$

where $IND_{IS}(B)$ is called the *B-indiscernibility* relation.

Observations

- An equivalence relation induces a partitioning of the universe.
- The partitions can be used to build new subsets of the universe.
- Subsets that are most often of interest have the same value of the decision attribute.
- It may happen, however, that a concept such as "Walk" cannot be defined in a crisp manner.

An Example of Set Approximation

U	A1	A2	Walk
x1	16-30	50	yes
x2	16-30	0	no
x3	31-45	1-25	no
x4	31-45	1-25	yes
x5	46-60	26-49	no
x6	16-30	26-49	yes
x7	46-60	26-49	no

- Let W = {x | Walk(x) = yes}.

$$\underline{A}W = \{x1, x6\},$$

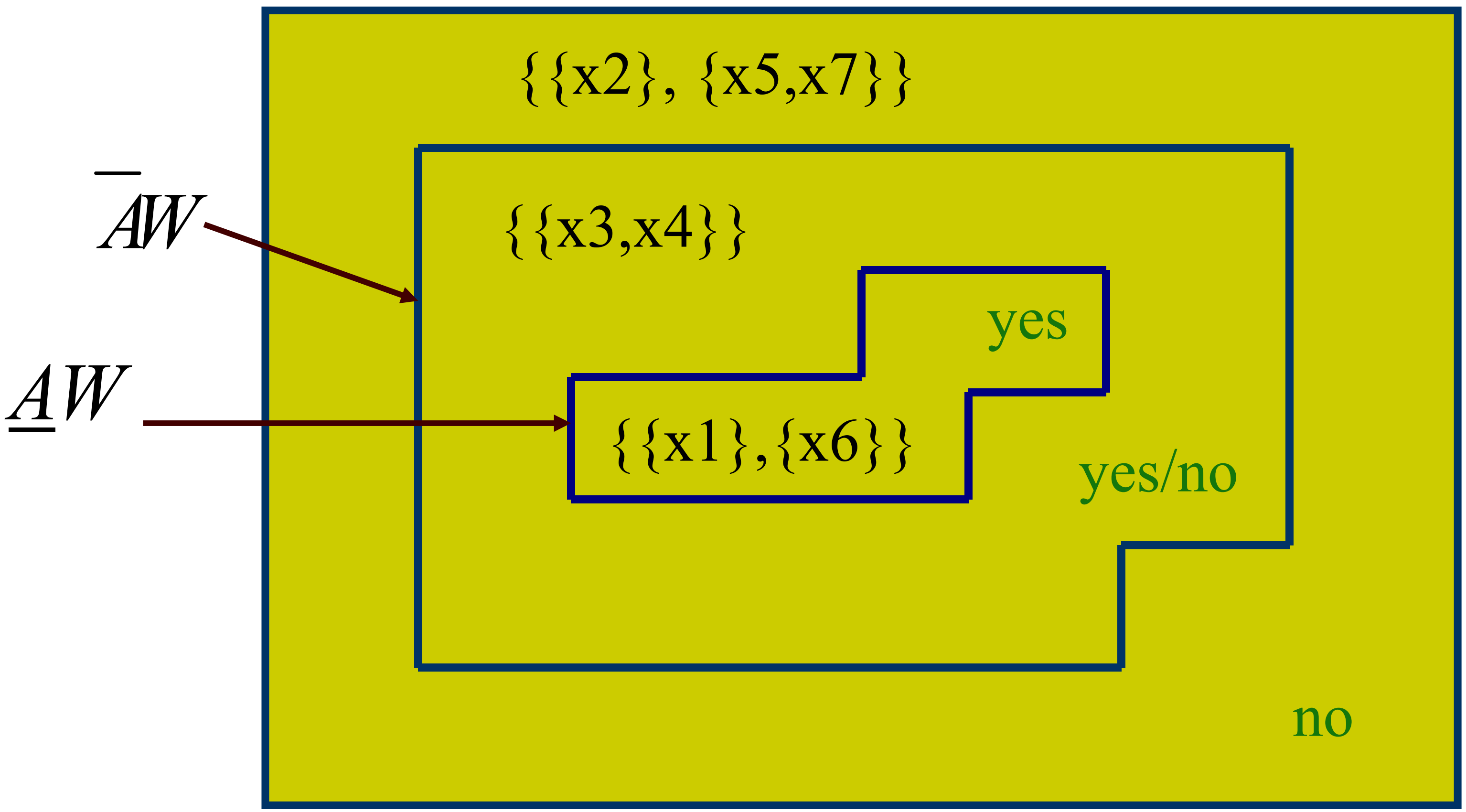
$$\overline{A}W = \{x1, x3, x4, x6\},$$

$$BN_A(W) = \{x3, x4\},$$

$$U - \overline{A}W = \{x2, x5, x7\}.$$

- The decision class, *Walk,* is rough since the boundary region is not empty.

An Example of Set Approximation



Set Approximation

- Let $T = (U, A)$ and let $B \subseteq A$ and $X \subseteq U$.

- We can approximate X using only the information contained in B by constructing the *B-lower* and *B-upper* approximations of X, denoted $\underline{B}X$ and $\overline{B}X$ respectively, where

$$\underline{B}X = \{x \mid [x]_B \subseteq X\},$$

$$\overline{B}X = \{x \mid [x]_B \cap X \neq \phi\}.$$

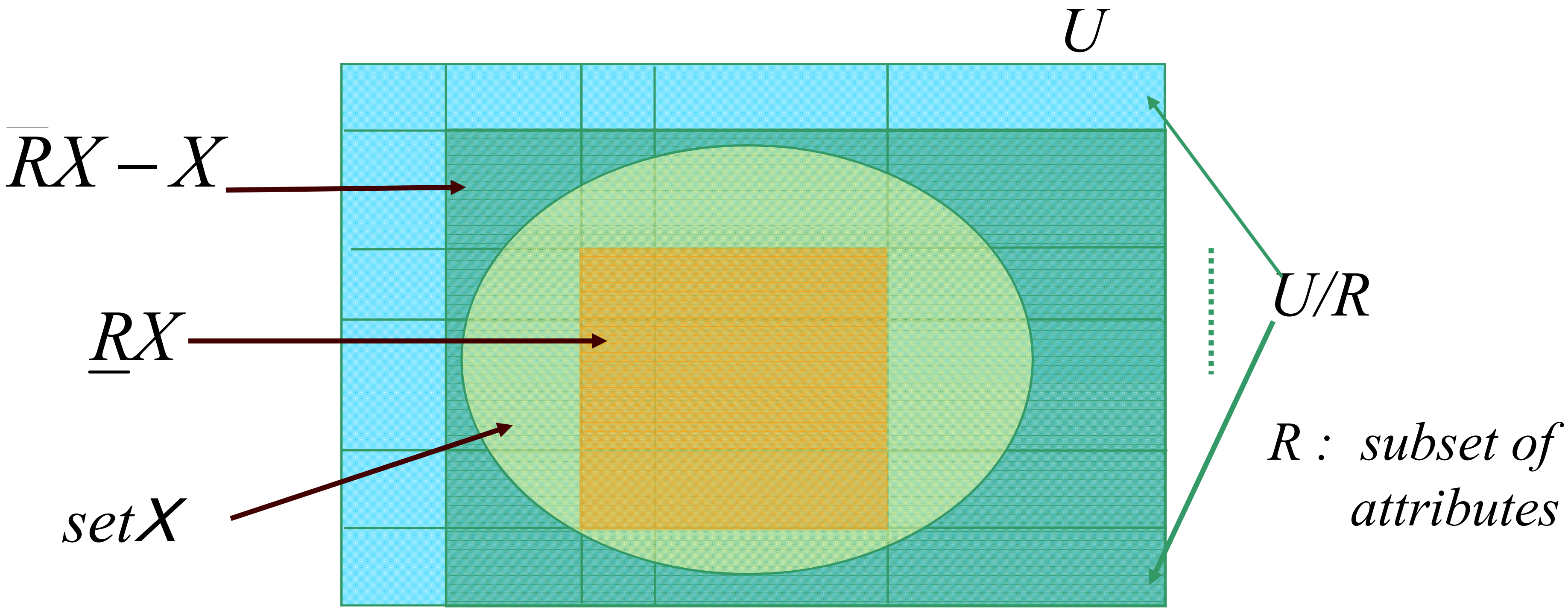
Set Approximation

- *B-boundary region* of *X*, $BN_B(X) = \overline{B}X - \underline{B}X$, consists of those objects that we cannot decisively classify into *X* in *B*.

- *B-outside region* of *X*, $U - \overline{B}X$, consists of those objects that can be with certainty classified as not belonging to *X*.

- A set is said to be *rough* if its boundary region is non-empty, otherwise the set is crisp.

Rough Sets



Properties of Approximations

$$\underline{B}(X) \subseteq X \subseteq \overline{B}X$$

$$\underline{B}(\phi) = \overline{B}(\phi) = \phi, \quad \underline{B}(U) = \overline{B}(U) = U$$

$$\overline{B}(X \cup Y) = \overline{B}(X) \cup \overline{B}(Y)$$

$$\underline{B}(X \cap Y) = \underline{B}(X) \cap \underline{B}(Y)$$

$X \subseteq Y$ implies $\underline{B}(X) \subseteq \underline{B}(Y)$ and $\overline{B}(X) \subseteq \overline{B}(Y)$

Properties of Approximations

$$\underline{B}(X \cup Y) \supseteq \underline{B}(X) \cup \underline{B}(Y)$$

$$\overline{B}(X \cap Y) \subseteq \overline{B}(X) \cap \overline{B}(Y)$$

$$\underline{B}(-X) = -\overline{B}(X)$$

$$\overline{B}(-X) = -\underline{B}(X)$$

$$\underline{B}(\underline{B}(X)) = \overline{B}(\underline{B}(X)) = \underline{B}(X)$$

$$\overline{B}(\overline{B}(X)) = \underline{B}(\overline{B}(X)) = \overline{B}(X)$$

where $-X$ denotes $U - X$.

Four Basic Classes of Rough Sets

- X is *roughly B-definable*, iff $\underline{B}(X) \neq \phi$ and $\overline{B}(X) \neq U$,

- X is *internally B-undefinable*, iff $\underline{B}(X) = \phi$ and $\overline{B}(X) \neq U$,

- X is *externally B-undefinable,* iff $\underline{B}(X) \neq \phi$ and $\overline{B}(X) = U$,

- X is *totally B-undefinable,* iff $\underline{B}(X) = \phi$ and $\overline{B}(X) = U$.

Accuracy of Approximation

$$\alpha_B(X) = \frac{|\underline{B}(X)|}{|\overline{B}(X)|}$$

where $|X|$ denotes the cardinality of $X \neq \phi$.

Obviously $0 \leq \alpha_B \leq 1$.

If $\alpha_B(X) = 1$, X is *crisp* with respect to *B*.

If $\alpha_B(X) < 1$, X is *rough* with respect to *B*.

Issues in the Decision Table/ Reducts

- Same or indiscernible objects may be represented several times.
- Some of the attributes may be superfluous (redundant). *That is, their removal cannot worsen the classification.*
- Keep only those attributes that preserve the indiscernibility relation and consequently set approximation.
- There are usually several such subsets of attributes and those which are minimal are called *reducts*.

Dispensable & Indispensable Attributes

- *C-positive region of D*: $POS_C(D) = \bigcup_{X \in U/D} \underline{C}X$

- Let $c \in C$. Attribute *c* is dispensable in *T* if $POS_C(D) = POS_{(C-\{c\})}(D)$, otherwise attribute *c* is indispensable in *T*.

- *T = (U, C, D)* is independent if all $c \in C$ are indispensable in *T*.

Reduct and Core

- The set of attributes $R \subseteq C$ is called a reduct of *C*, if *T'* = (*U*, *R*, *D*) is independent and

$$POS_R(D) = POS_C(D).$$

- The set of all the condition attributes indispensable in T is denoted by *CORE(C)*.

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

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

An Example of Reducts and Core

U	*Headache*	*Muscle pain*	*Temp.*	*Flu*
U1	Yes	Yes	Normal	No
U2	Yes	Yes	High	Yes
U3	Yes	Yes	Very-high	Yes
U4	No	Yes	Normal	No
U5	No	No	High	No
U6	No	Yes	Very-high	Yes

Reduct1 **= {Muscle-pain,Temp.}**

U	*Muscle pain*	*Temp.*	*Flu*
U1,U4	Yes	Normal	No
U2	Yes	High	Yes
U3,U6	Yes	Very-high	Yes
U5	No	High	No

Reduct2 **= {Headache, Temp.}**

U	*Headache*	*Temp.*	*Flu*
U1	Yes	Norlmal	No
U2	Yes	High	Yes
U3	Yes	Very-high	Yes
U4	No	Normal	No
U5	No	High	No
U6	No	Very-high	Yes

$$\textit{CORE} = \{\text{Headache,Temp}\} \cap \{\text{MusclePain, Temp}\} = \{\text{Temp}\}$$

Dependency of Attributes

- Discovering dependencies between attributes is an important issue in KDD.

- A set of attribute D depends totally on a set of attributes C, denoted $C \Rightarrow D$, if all values of attributes from D are uniquely determined by values of attributes from C.

Dependency of Attributes

- Let D and C be subsets of A. We will say that D depends on C in a degree k $(0 \leq k \leq 1)$,

denoted by $C \Rightarrow_k D$, if

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

where $POS_C(D) = \bigcup_{X \in U/D} \underline{C}(X)$, called C-positive region of D.

Dependency of Attributes

- Obviously

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

- 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 Membership

- The rough membership function quantifies the degree of relative overlap between the set X and the equivalence class $[x]_B$ to which x belongs.

$$\mu_X^B : U \to [0,1] \qquad \mu_X^B = \frac{|[x]_B \cap X|}{|[x]_B|}$$

- The rough membership function can be interpreted as a frequency-based estimate of $P(x \in X \mid u)$, where u is the equivalence class of $IND(B)$.

Rough Membership

- The formulae for the lower and upper approximations can be generalized to some arbitrary level of precision $\pi \in (0.5, 1]$ by means of the rough membership function

$$\underline{B}_{\pi} X = \{x \mid \mu_X^B(x) \geq \pi\}$$

$$\overline{B}_{\pi} X = \{x \mid \mu_X^B(x) > 1 - \pi\}.$$

- Note: the lower and upper approximations as originally formulated are obtained as a special case with $\pi = 1$.

Feature Selection

- In real-data analysis, data set may contain a number of redundant and insignificant features with low relevance to the classes.
  - The presence of such redundant, insignificant, and non-relevant features leads to a reduction in the useful information.
- Ideally, the selected features should have high significance and relevance with classes, while redundancy among them would be as low as possible.
  - Expected to be able to predict the classes of the samples.
- Hence, to assess the effectiveness of the features, relevance, significance, and redundancy need to be measured quantitatively.
- Rough sets, entropy, mutual information, *f*-information

Max-Dependency Criterion

- Max-Dependency: Select a subset of features (condition attributes) which jointly have the largest dependency on the target class (decision attribute)

$$\max \mathscr{D}(\mathbb{S},\mathbb{D}),\ \mathscr{D} = \gamma_{\{\mathscr{A}_i, i=1,\cdots,d\}}(\mathbb{D}),$$

degree of dependency

$$\boxed{\gamma_{\mathbb{C}}(\mathbb{D}) = \frac{|POS_{\mathbb{C}}(\mathbb{D})|}{|\mathbb{U}|}}$$

- Equivalently, $\max\limits_{\mathscr{A}_j \in \{\mathbb{C}-\mathbb{S}_{d-1}\}} \{\gamma_{\{\mathbb{S}_{d-1},\mathscr{A}_j\}}(\mathbb{D}) - \gamma_{\mathbb{S}_{d-1}}(\mathbb{D})\} = \max\limits_{\mathscr{A}_j \in \{\mathbb{C}-\mathbb{S}_{d-1}\}} \{\sigma_{\mathbb{S}_d}(\mathbb{D},\mathscr{A}_j)\}$

significance

- Problem of Max-Dependency for real life applications:
  - Not sufficient for selecting highly discriminative features
  - hard to generate resultant equivalence classes in the high-dimensional space: the number of samples is often insufficient and the generation of resultant equivalence classes is usually an ill-posed problem.
  - slow computational speed

Max-Relevance-Max-Significance Criterion

- Max-Relevance: Select a subset of features which approximates Max-Dependency with the mean value of all dependency values between individual feature and target class label $\max \mathscr{R}(\mathbb{S},\mathbb{D}), \qquad \mathscr{R} = \frac{1}{|\mathbb{S}|}\sum_{\mathscr{A}_i \in \mathbb{S}} \gamma_{\mathscr{A}_i}(\mathbb{D}).$
  - Features selected according to Max-Relevance could have rich redundancy, that is, dependency among them could be large.
  - When two features highly depend on each other, the respective class discriminative power would not change much if one of them were removed.
- Max-Significance: Select mutually exclusive features

$$\max \mathscr{S}(\mathbb{S},\mathbb{D}),\ \mathscr{S} = \frac{1}{|\mathbb{S}|(|\mathbb{S}|-1)} \sum_{\mathscr{A}_i \neq \mathscr{A}_j \in \mathbb{S}} \{\sigma_{\{\mathscr{A}_i,\mathscr{A}_j\}}(\mathbb{D},\mathscr{A}_i) + \sigma_{\{\mathscr{A}_i,\mathscr{A}_j\}}(\mathbb{D},\mathscr{A}_j)\}$$

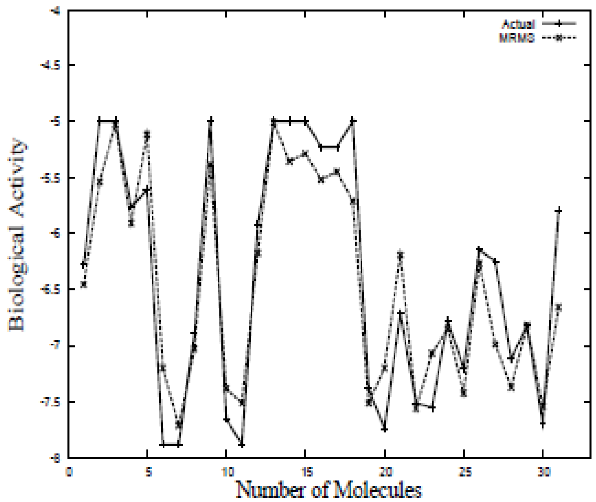
$$\max \Phi(\mathscr{R},\mathscr{S}), \qquad \Phi = \mathscr{R} + \mathscr{S}.$$

- Max-Relevance-Max-Significance Criterion:

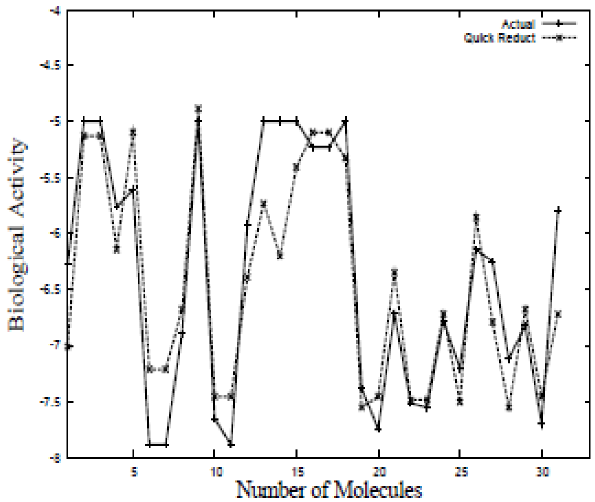
P. Maji and S. Paul, "Rough Set Based Maximum Relevance-Maximum Significance Criterion and Gene Selection from Microarray Data", *International Journal of Approximate Reasoning*, 52(3), pp. 408--426, March 2011.

Prediction of Biological Activity of Molecules

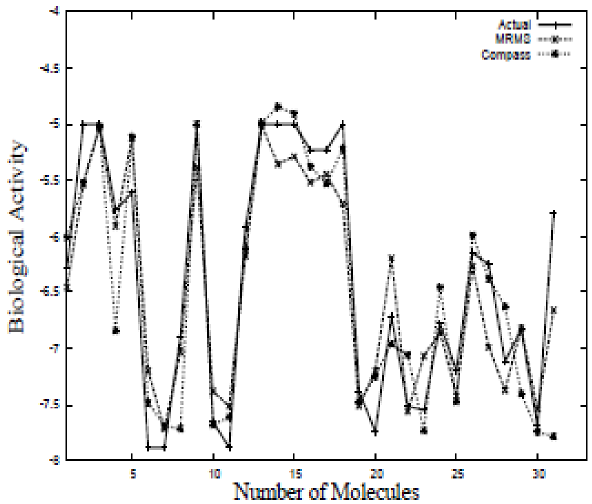
Actual / MRMS



Actual / MD



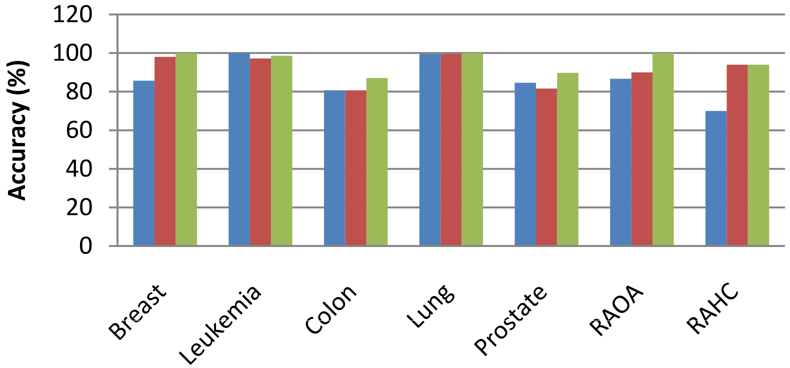
Actual / MRMS / Compass

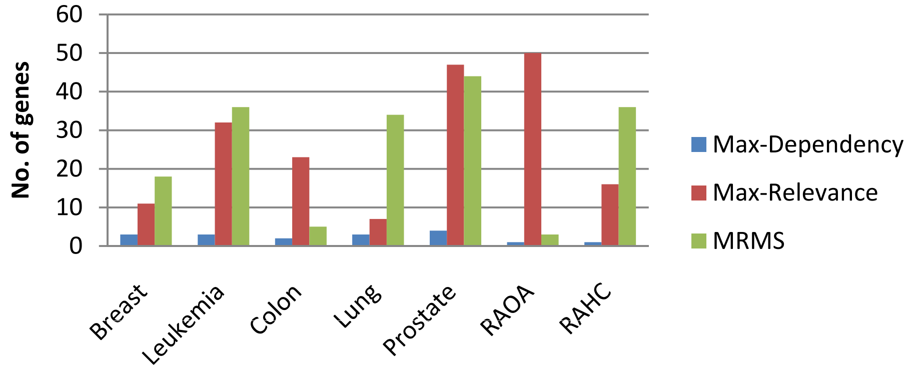


- Steroid molecules: 31 samples and 4000 molecular descriptors
- MRMS / MD / Compass:
  - R^2 statistic of support vector regression method: 0.88 / 0.82 / 0.79
  - execution time: 3498 ms / 383253 ms / 111600 ms

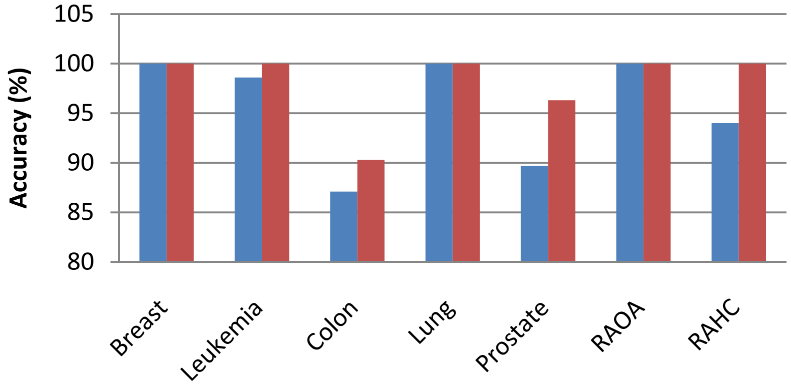
P. Maji and S. Paul, ``Rough Sets for Selection of Molecular Descriptors to Predict Biological Activity of Molecules'', *IEEE Trans. Systems, Man and Cybernetics, Part C, Applications and Reviews, 40*(6), pp. 639--648, 2010.

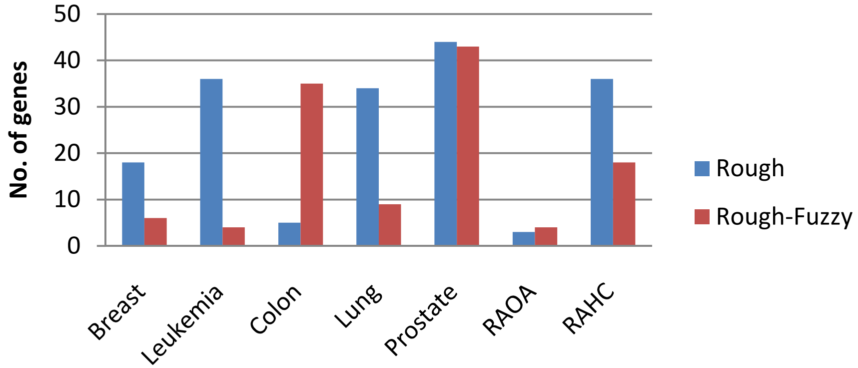
Gene Selection from Microarray Data





MRMS provides higher classification accuracy of SVM with lower number of genes





Rough-fuzzy provides higher classification accuracy of SVM with lower number of genes than rough

Fuzzy Equivalence Partition Matrix

If c and n denote the number of fuzzy information granules (equivalence classes) and number of objects in U, then c-partitions of U can be conveniently arrayed as a ($c \times n$) fuzzy equivalence partition matrix

$$\mathbb{M}_{\mathbb{A}} = \begin{pmatrix} m_{11}^{\mathbb{A}} & m_{12}^{\mathbb{A}} & \cdots & m_{1n}^{\mathbb{A}} \\ m_{21}^{\mathbb{A}} & m_{22}^{\mathbb{A}} & \cdots & m_{2n}^{\mathbb{A}} \\ \cdots & \cdots & \cdots & \cdots \\ m_{c1}^{\mathbb{A}} & m_{c2}^{\mathbb{A}} & \cdots & m_{cn}^{\mathbb{A}} \end{pmatrix}$$

where m_{ij}^{A} Є [0, 1] represents membership of object x_j in ith fuzzy equivalence partition F_i

Fuzzy relative frequency corresponding to fuzzy equivalence partition F_i is

$$\lambda_{F_i} = \frac{1}{n}\sum_{j=1}^{n} m_{ij}^{\mathbb{A}},$$

If p and q are the number of fuzzy equivalence partitions generated by fuzzy attribute sets P and Q, and P_i and Q_j represent the corresponding ith and jth fuzzy equivalence partitions, respectively, then, joint frequency of P_i and Q_j is

$$\lambda_{P_i Q_j} = \frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}} \cap m_{jk}^{\mathbb{Q}}).$$

f-Information Measures

- Entropy (on fuzzy approximation spaces of fuzzy attribute set ***A***):

$$H(\mathbb{A}) = -\sum_{i=1}^{c}\left[\frac{1}{n}\sum_{k=1}^{n} m_{ik}^{\mathbb{A}}\right]\log\left[\frac{1}{n}\sum_{k=1}^{n} m_{ik}^{\mathbb{A}}\right].$$

- Mutual information (between two fuzzy attribute sets ***P*** and ***Q***):

$$I(\mathbb{P},\mathbb{Q}) = -\sum_{i=1}^{p}\left[\frac{1}{n}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\right]\log\left[\frac{1}{n}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\right] - \sum_{j=1}^{q}\left[\frac{1}{n}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right]\log\left[\frac{1}{n}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right] + \sum_{i=1}^{p}\sum_{j=1}^{q}\left[\frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}}\cap m_{jk}^{\mathbb{Q}})\right]\log\left[\frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}}\cap m_{jk}^{\mathbb{Q}})\right]$$

- Some f-information measures between two fuzzy attribute sets ***P*** and ***Q***

$$V(\mathbb{P},\mathbb{Q}) = \sum_{i=1}^{p}\sum_{j=1}^{q}\left|\frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}}\cap m_{jk}^{\mathbb{Q}}) - \frac{1}{n^2}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right|.$$

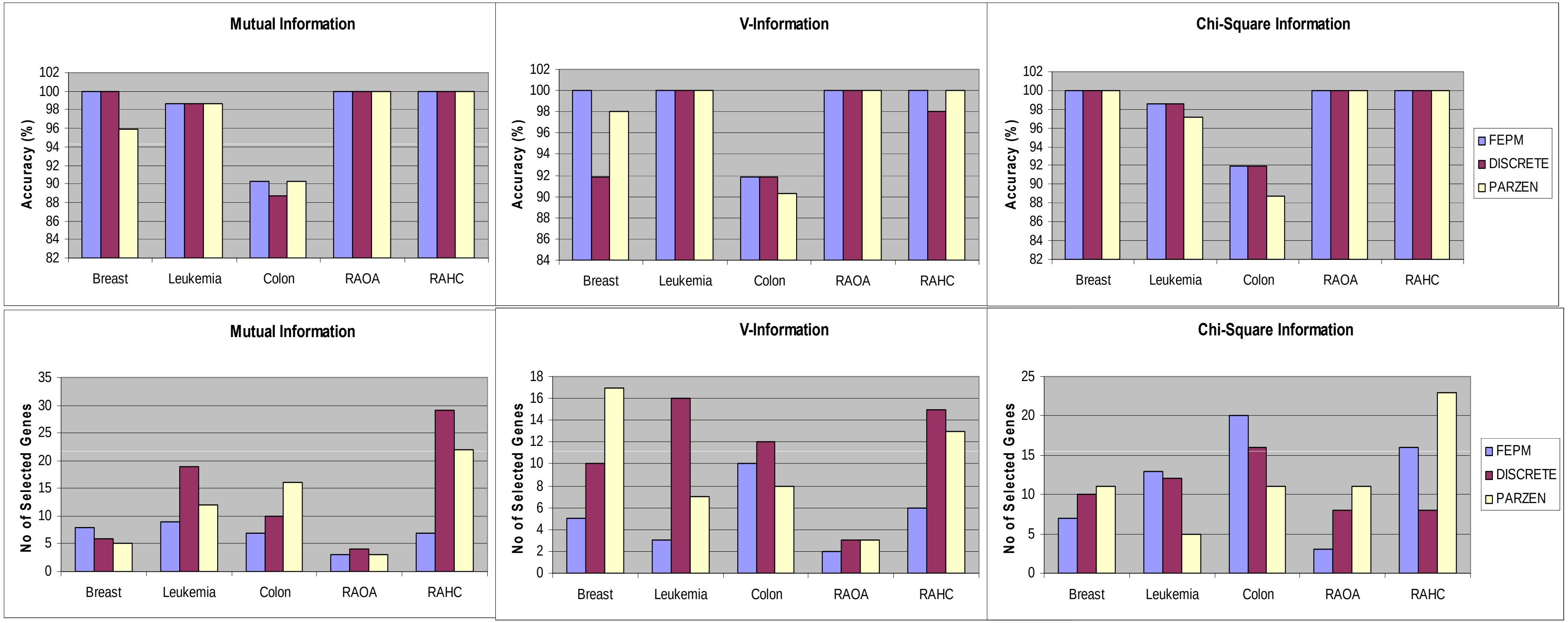
$$M_\alpha(\mathbb{P},\mathbb{Q}) = \sum_{i=1}^{p}\sum_{j=1}^{q}\left|\left[\frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}}\cap m_{jk}^{\mathbb{Q}})\right]^{\alpha} - \left[\frac{1}{n^2}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right]^{\alpha}\right|^{\frac{1}{\alpha}}$$

$$I_\alpha(\mathbb{P},\mathbb{Q}) = \frac{1}{\alpha(\alpha-1)}\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\frac{\left[\frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}}\cap m_{jk}^{\mathbb{Q}})\right]^{\alpha}}{\left[\frac{1}{n^2}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right]^{\alpha-1}} - 1\right)$$

$$\chi^\alpha(\mathbb{P},\mathbb{Q}) = \sum_{i=1}^{p}\sum_{j=1}^{q}\frac{\left|\frac{1}{n}\sum_{k=1}^{n}(m_{ik}^{\mathbb{P}}\cap m_{jk}^{\mathbb{Q}}) - \frac{1}{n^2}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right|^{\alpha}}{\left(\frac{1}{n^2}\sum_{k=1}^{n} m_{ik}^{\mathbb{P}}\sum_{k=1}^{n} m_{jk}^{\mathbb{Q}}\right)^{\alpha-1}}.$$

P. Maji and S. K. Pal, ``Feature Selection Using f-Information Measures in Fuzzy Approximation Spaces'', *IEEE Trans. Systems, Man and Cybernetics, Part B, Cybernetics,* 22(6), pp. 854--867, 2010.

Gene Selection from Microarray Data



- Information measures are used to compute both relevance and redundancy
  - Maximization of relevance and minimization of redundancy

Higher or same accuracy with lower number of genes in case of FEPM approach

P. Maji and S. K. Pal, ``Fuzzy-Rough Sets for Information Measures and Selection of Relevant Genes from Microarray Data'', *IEEE Trans. Systems, Man and Cybernetics, Part B, Cybernetics,* 40(3), pp. 741--752, 2010.

Outline

- **Rough-Fuzzy Computing**

- **Clustering: Hard and Fuzzy**

- **Rough-Fuzzy Clustering**
  - **Design of Algorithm**
  - **Objective Function; Cluster Prototypes; Membership Functions**
  - **Convergence of Algorithm; Generalization of Existing Algorithms**

- **Performance Analysis**
  - **Numerical Data Sets (Clustering Benchmark Data; Segmentation of Brain MR Images; Text-Graphics Segmentation; Clustering Functionally Similar Genes)**
  - **Non-numerical Data Sets (Amino Acid Sequence Analysis)**

- **Conclusion**

Rough Sets

- It approximates an arbitrary set X by a pair of lower and upper approximations
  - Lower approximation $\underline{R}(X)$: collection of those elements of U (the universe of discourse) that definitely belong to X;
  - Upper approximation $\bar{R}(X)$: collection of those elements of U that possibly belong to X.
  - The interval $[\underline{R}(X), \bar{R}(X)]$ is the representation of an ordinary set X in the approximation space $< U, R >$ or simply called the rough set of X.

- Key notions of rough set:
  - Information granule (clump of similar objects): formalizes the concept of finite precision representation of objects in real life situation; and
  - Reducts: represent the core of an information system, both in terms of objects and features, in a granular universe.

Granular Computing

- Granular computing refers to that domain where computation and operations are performed on information granules.

- Granular computing is applicable to the situations
  1) when a problem involves incomplete, uncertain and vague information, it may be difficult to differentiate distinct elements and one may find it convenient to consider granules for its handling
  2) though detailed information is available, it may be sufficient to use granules in order to have an efficient and practical solution.

- The computation time is greatly reduced as computations are performed on granules, rather than on individual data points.
  - useful for designing scalable pattern recognition algorithms

Rough-Fuzzy Computing

- There are usually real valued data and fuzzy information in real world applications.

- Fuzzy set theory hinges on the notion of a membership function on the domain of discourse, assigning to each object a grade of belongingness in order to represent an imprecise or overlapping concept.

- Rough-Fuzzy Computing: Combining fuzzy sets and rough sets provides a mathematical framework to capture uncertainties associated with the data. They are complementary in some aspects:
  - while the membership functions of fuzzy sets enable efficient handling of overlapping classes,
  - the concept of lower and upper approximations of rough sets deals with uncertainty, vagueness, and incompleteness in class definitions.

Clustering

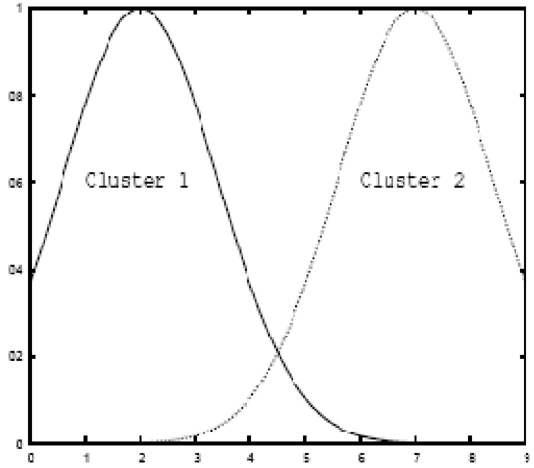
- Cluster Analysis: a technique for finding natural groups present in data
  - Divides a data set into a set of clusters in such a way that two objects from the same cluster are as similar as possible and the objects from different clusters are as dissimilar as possible.

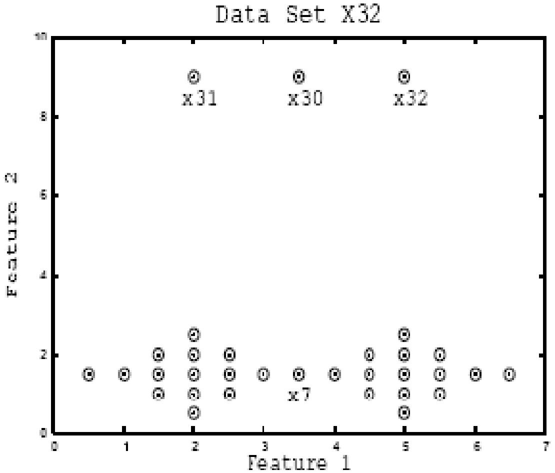
- Hard Clustering: k-means or hard c-means (HCM) algorithm

- Basic steps of HCM:
  - Assign initial means or centroids v_i, $i = 1, 2, \cdots, c$.
  - For each object x_j, calculate distance d_{ij} between itself and the centroid v_i of ith cluster.
  - If d_{ij} is minimum for $1 <= i <= c$, then assign x_j to ith cluster.
  - Compute new centroids.
  - Repeat above three steps until no more new assignments can be made.

Hard Clustering

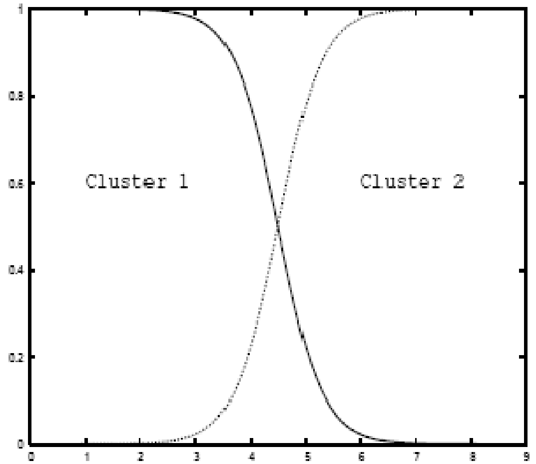
- Hard Clustering:
  - each object is assigned to exactly one cluster.
  - In real data analysis, uncertainties arise due to overlapping cluster boundaries and incompleteness and vagueness in cluster definition.
  - Also, noise and outliers are unavoidable.

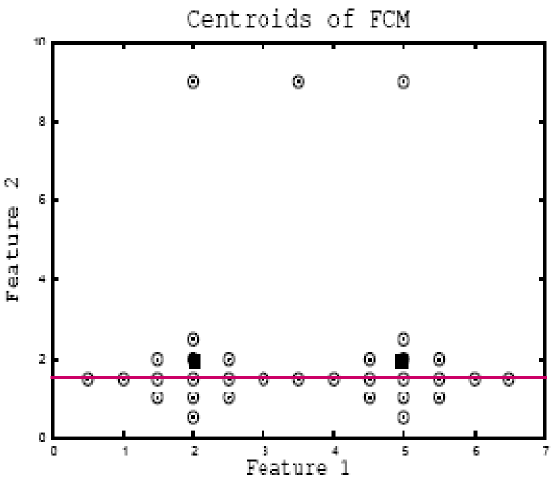




Fuzzy Clustering

- **Fuzzy *c*-means (FCM)**: allows gradual memberships - assigns memberships to an object which are inversely related to the relative distance of the object to the cluster prototypes.
- Offers the opportunity to deal with the data that belong to more than one cluster at the same time
  - Can deal - uncertainties arising from overlapping cluster boundaries
- Resulting membership values of FCM
  - Do not always correspond well to degrees of belonging of data
  - May be inaccurate in a noisy environment.

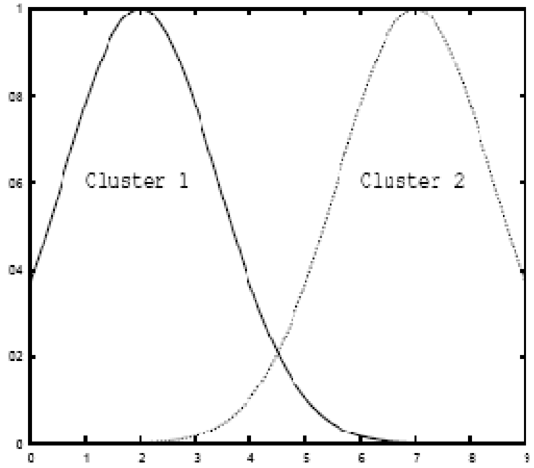




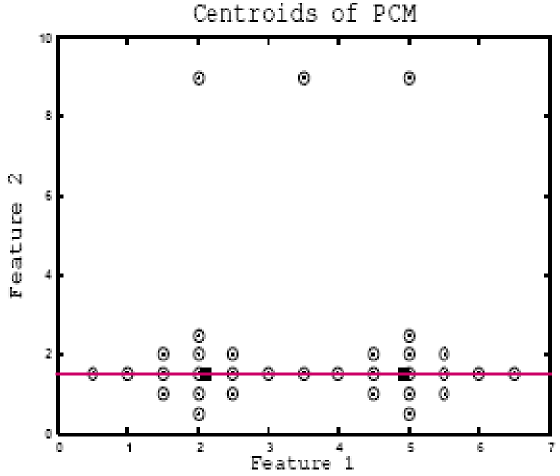
J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithm*. New York: Plenum, 1981.

Possibilistic Clustering

- Possibilistic c-means (PCM): uses a possibilistic type of membership function
  - Produces memberships that have a good explanation of degrees of belonging for data
  - Reduces weakness of FCM in noisy environment



- Resulting membership values of PCM
  - may generate coincident clusters

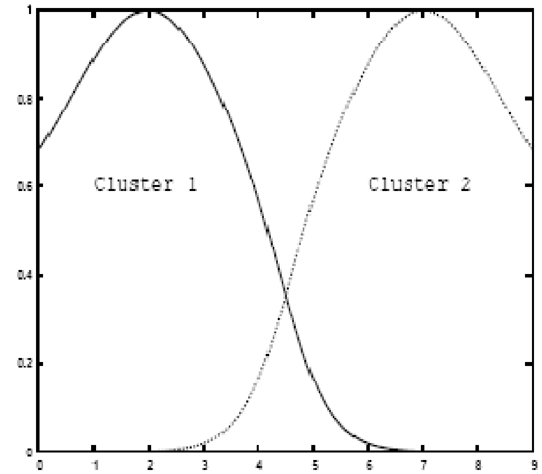


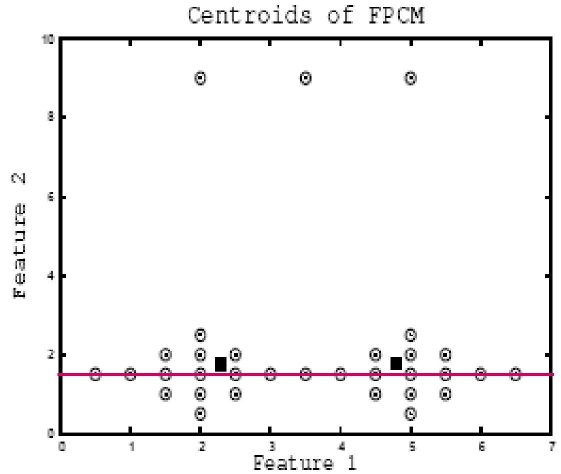
R. Krishnapuram and J. M. Keller, "A Possibilistic Approach to Clustering," *IEEE Trans. Fuzzy Systems*, vol. 1, no. 2, pp. 98–110, 1993.

Fuzzy-Possibilistic Clustering

- Fuzzy-possibilistic *c*-means (FPCM): Uses both fuzzy and possibilistic memberships

  - Weighted average of FCM and PCM
  - Reduces weakness of FCM in noisy environment
  - Reduces coincident clusters generation problem of PCM





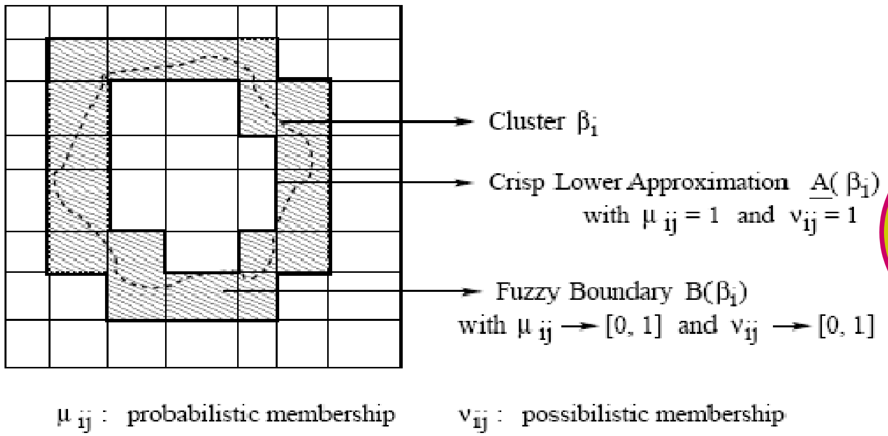
N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A Possibilistic Fuzzy C-Means Clustering Algorithm," *IEEE Trans. Fuzzy Systems*, vol. 13, no. 4, pp. 517–530, 2005.

F. Masulli and S. Rovetta, "Soft Transition from Probabilistic to Possibilistic Fuzzy Clustering," *IEEE Trans. Fuzzy Systems*, vol. 14, no. 4, pp. 516–527, 2006.

Rough-Fuzzy Clustering

- **Rough-Fuzzy-Possibilistic Clustering**: Adds memberships of fuzzy sets, and lower and upper approximations of rough sets into hard clustering.

- While membership function enables handling of overlapping partitions, rough sets deal with vagueness and incompleteness in class



Each cluster: represented by a cluster prototype, a crisp lower approximations, and a fuzzy (probabilistic and possibilistic) boundary

P. Maji and S. K. Pal, ``Rough Set Based Generalized Fuzzy *C*-Means Algorithm and Quantitative Indices'', *IEEE Trans. System, Man and Cybernetics, Part B, Cybernetics*, 37(6), pp. 1529--1540, 2007.

Rough-Fuzzy-Possibilistic Clustering

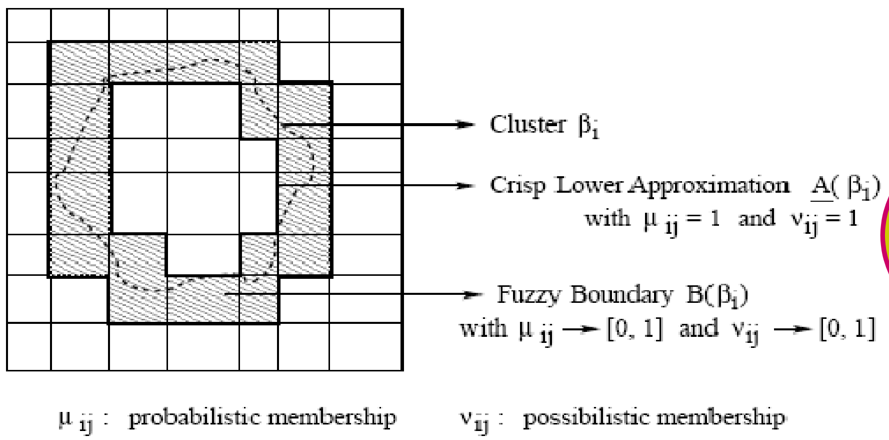
- According to rough sets, if x_j belongs to lower approximation of ith cluster, it does not belong to lower approximation of any other clusters - that is, x_j is contained in ith cluster definitely
- Hence, the memberships of objects in lower approximation of a cluster
  - should be independent of other centroids and clusters, and
  - should not be coupled with their similarity with respect to other cluster prototypes.
  - Also, objects in lower approximation of a cluster should have similar influence on corresponding cluster prototype and cluster.
- If x_j belongs to boundary region of ith cluster, it possibly belongs to ith cluster and potentially belongs to other clusters - hence, objects in boundary regions should have different influence on cluster prototypes and clusters.

P. Maji and S. K. Pal, ``Rough Set Based Generalized Fuzzy *C*-Means Algorithm and Quantitative Indices'', *IEEE Trans. System, Man and Cybernetics, Part B, Cybernetics*, 37(6), pp. 1529--1540, 2007.

Rough-Fuzzy-Possibilistic Clustering

- So, membership values of objects in lower approximation are 1 while those in boundary region are same as fuzzy and possibilistic clustering.

- It partitions each cluster into two classes: lower approximation (core) and boundary (overlapping). Only objects in boundary are fuzzified.



Each cluster: represented by a cluster prototype, a crisp lower approximations, and a fuzzy (probabilistic and possibilistic) boundary

P. Maji and S. K. Pal, ``Rough Set Based Generalized Fuzzy *C*-Means Algorithm and Quantitative Indices'', *IEEE Trans. System, Man and Cybernetics, Part B, Cybernetics*, 37(6), pp. 1529--1540, 2007.

Objective Function

$$J_{\text{RFP}} = \begin{cases} w \times \mathcal{A}_1 + \tilde{w} \times \mathcal{B}_1 & \text{if } \underline{A}(\beta_i) \neq \emptyset,\ B(\beta_i) \neq \emptyset \\ \mathcal{A}_1 & \text{if } \underline{A}(\beta_i) \neq \emptyset,\ B(\beta_i) = \emptyset \\ \mathcal{B}_1 & \text{if } \underline{A}(\beta_i) = \emptyset,\ B(\beta_i) \neq \emptyset \end{cases}$$

$w + \tilde{w} = 1;$ $a + b = 1.$
$0 < \tilde{w} < w < 1,$
$0 < a < 1,$ and $0 < b < 1,$

$$\mathcal{A}_1 = \sum_{i=1}^{c} \sum_{x_j \in \underline{A}(\beta_i)} \{a(\mu_{ij})^{\acute{m}_1} + b(\nu_{ij})^{\acute{m}_2}\}||x_j - v_i||^2 + \sum_{i=1}^{c} \eta_i \sum_{x_j \in \underline{A}(\beta_i)} (1-\nu_{ij})^{\acute{m}_2} = \sum_{i=1}^{c} \sum_{x_j \in \underline{A}(\beta_i)} ||x_j - v_i||^2$$

$$\mathcal{B}_1 = \sum_{i=1}^{c} \sum_{x_j \in B(\beta_i)} \{a(\mu_{ij})^{\acute{m}_1} + b(\nu_{ij})^{\acute{m}_2}\}||x_j - v_i||^2 + \sum_{i=1}^{c} \eta_i \sum_{x_j \in B(\beta_i)} (1-\nu_{ij})^{\acute{m}_2}$$

- x_j and v_i: object and cluster prototype or centroid
- w and $\hat{w}$ $(= 1-w)$: relative importance of lower and boundary region
- a and b: relative importance of probabilistic and possibilistic memberships
- μ_{ij} and v_{ij} : probabilistic and possibilistic memberships
- m_1 and m_2: probabilistic and possibilistic fuzzifiers
- η_i : scale parameter, represents the size of ith cluster

Cluster Prototype and Memberships

$$v_i^{\mathrm{RFP}} = \begin{cases} w \times \mathcal{C}_1 + \tilde{w} \times \mathcal{D}_1 & \text{if } A(\beta_i) \neq \emptyset,\ B(\beta_i) \neq \emptyset \\ \mathcal{C}_1 & \text{if } \underline{A}(\beta_i) \neq \emptyset,\ B(\beta_i) = \emptyset \\ \mathcal{D}_1 & \text{if } \underline{A}(\beta_i) = \emptyset,\ B(\beta_i) \neq \emptyset \end{cases}$$

$w + \tilde{w} = 1;$ $a + b = 1.$
$0 < \tilde{w} < w < 1,$
$0 < a < 1,$ and $0 < b < 1,$

$$\mathcal{C}_1 = \frac{1}{|\underline{A}(\beta_i)|} \sum_{x_j \in \underline{A}(\beta_i)} x_j$$ where $|\underline{A}(\beta_i)|$ → ***Cardinality of lower approximation***

$$\mathcal{D}_1 = \frac{1}{n_i} \sum_{x_j \in B(\beta_i)} \{a(\mu_{ij})^{\acute{m}_1} + b(\nu_{ij})^{\acute{m}_2}\} x_j \quad \text{where} \quad n_i = \sum_{x_j \in B(\beta_i)} \{a(\mu_{ij})^{\acute{m}_1} + b(\nu_{ij})^{\acute{m}_2}\}$$

→ ***Cardinality of boundary region***

$$\mu_{ij} = \left(\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{\acute{m}_1 - 1}}\right)^{-1}; \text{ where } d_{ij}^2 = ||x_j - v_i||^2$$

$$\nu_{ij} = \frac{1}{1 + \mathrm{E}}; \quad \text{where} \quad \mathrm{E} = \left\{\frac{b||x_j - v_i||^2}{\eta_i}\right\}^{1/(\acute{m}_2 - 1)}$$

$$\eta_i = \frac{\sum_{j=1}^{n} (\nu_{ij})^{\acute{m}_2} ||x_j - v_i||^2}{\sum_{j=1}^{n} (\nu_{ij})^{\acute{m}_2}}$$

Convergence of Proposed Algorithm

- **Cluster prototype:** $v_i^{RFP} = w \times \mathrm{v}_i^{RFP} + \hat{w} \times \acute{\upsilon}_i^{RFP}$
- **Convergence of v_i^{RFP} depends on convergence of v_i^{RFP} and $\acute{\upsilon}_i^{RFP}$**

$$(|\underline{A}(\beta_i)|)\mathrm{v}_i^{\mathrm{RFP}} = \sum_{x_j \in \underline{A}(\beta_i)} x_j$$

$$(n_i)\tilde{\mathrm{v}}_i^{\mathrm{RFP}} = \sum_{x_j \in B(\beta_i)} \{a(\mu_{ij})^{\acute{m}_1} + b(\nu_{ij})^{\acute{m}_2}\}x_j$$

A set of linear equations in terms of v_i^{RFP} and $\acute{\upsilon}_i^{RFP}$ if both μ_{ij} and v_{ij} are kept constant

Gauss-Seidel Algorithm

$$\mathrm{A} = \begin{bmatrix} |\underline{A}(\beta_1)| & 0 & \cdots & \cdots & 0 \\ 0 & |\underline{A}(\beta_2)| & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & |\underline{A}(\beta_c)| \end{bmatrix} \quad \tilde{\mathrm{A}} = \begin{bmatrix} n_1 & 0 & \cdots & \cdots & 0 \\ 0 & n_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & n_c \end{bmatrix}$$

Both Matrices to be diagonally dominant, we must have

$$|\underline{A}(\beta_i)| > 0; \quad \text{and} \quad n_i > 0.$$

sufficient condition, not necessary

Bezdek's Theorem

If the cluster prototypes and memberships are computed alternatively in iteration, then the proposed algorithm converges, at least along a subsequence, to a local optimum solution.

Core and Boundary Region

- It starts by choosing c objects as centroids of c clusters. The fuzzy and possibilistic memberships of all objects are then calculated.

- After computing the memberships for c clusters and n objects, difference of two highest memberships of each object is compared with a threshold value δ.

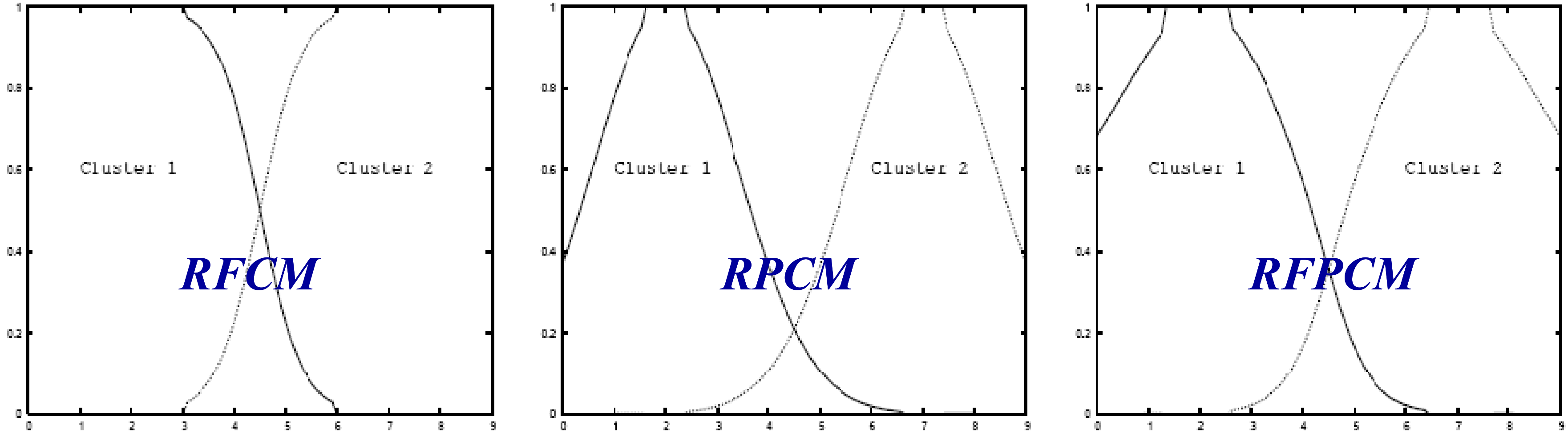
- Let u_{ij} and u_{kj} be the highest and second highest memberships of object x_j, where $u_{ij} = a\ \mu_{ij} + b\ v_{ij}$.
  - If $(u_{ij} - u_{kj}) < \delta$, then x_j belongs to boundary regions of ith and kth clusters and x_j does not belong to lower approximation of any cluster
  - Otherwise x_j belongs to lower approximation (core) of ith cluster.

- After defining lower approximations and boundary regions of different clusters based on δ, the memberships u_{ij} of objects in lower approximation are set to 1, while those in boundary regions are remain unchanged.

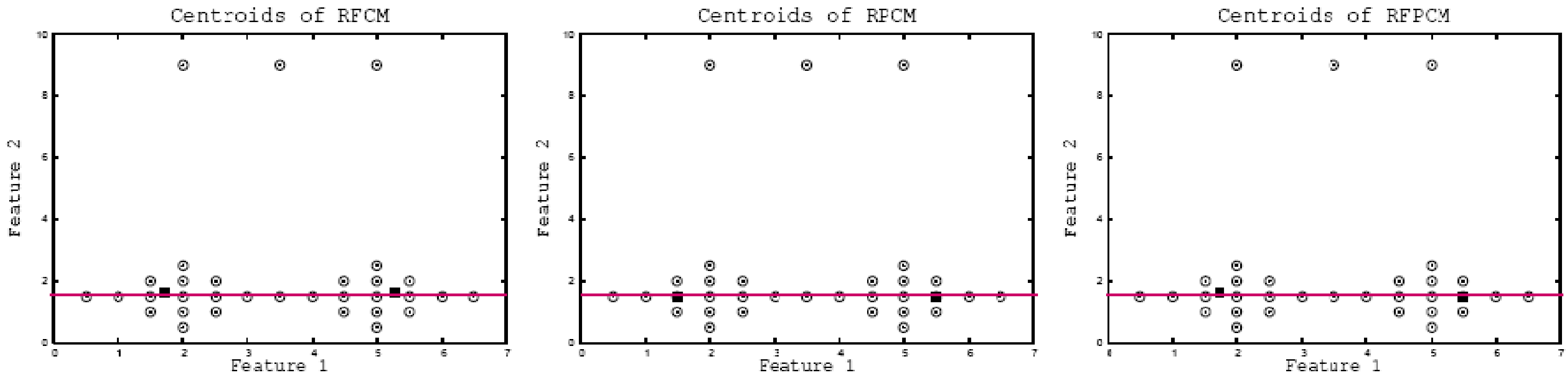
Generalization of Existing Algorithms

- If $\delta = 0$, RFPCM reduces to HCM (hard c-means).
- If $\delta = 1$, RFPCM reduces to FPCM (fuzzy-possibilistic c-means).
  - $a = 1$, RFPCM reduces to FCM (fuzzy c-means).
  - $a = 0$, RFPCM reduces to PCM (possibilistic c-means).
- If $\delta \neq 0, 1$ and $a = 1$, RFPCM reduces to RFCM (rough-fuzzy c-means).
- If $\delta \neq 0, 1$ and $a = 0$, RFPCM reduces to RPCM (rough-possibilistic c-means).
- The value of threshold, $\delta = \frac{1}{n}\sum_{j=1}^{n}(u_{ij} - u_{kj})$ ➡ Represents average difference of two highest memberships of all objects

Different Rough-Fuzzy Clustering



For objects near to centers and far from separating line, different shape of memberships is apparent.



Quantitative Indices

Davies-Bouldin (DB) index is a function of the ratio of sum of within-cluster distance to between-cluster separation:

$$\mathrm{DB} = \frac{1}{c}\sum_{i=1}^{c} \max_{i \neq k}\left\{\frac{S(v_i) + S(v_k)}{d(v_i, v_k)}\right\}$$

A good clustering procedure should make DB index value as low as possible.

Dunn's (D) index is designed to identify sets of clusters that are compact and well separated. It maximizes

$$D = \min_i\left\{\min_{i \neq k}\left\{\frac{d(v_i, v_k)}{\max_l S(v_l)}\right\}\right\}$$

A good clustering procedure should make D- index value as high as possible.

β-Index: It is defined as ratio of total variation and within-cluster variation:

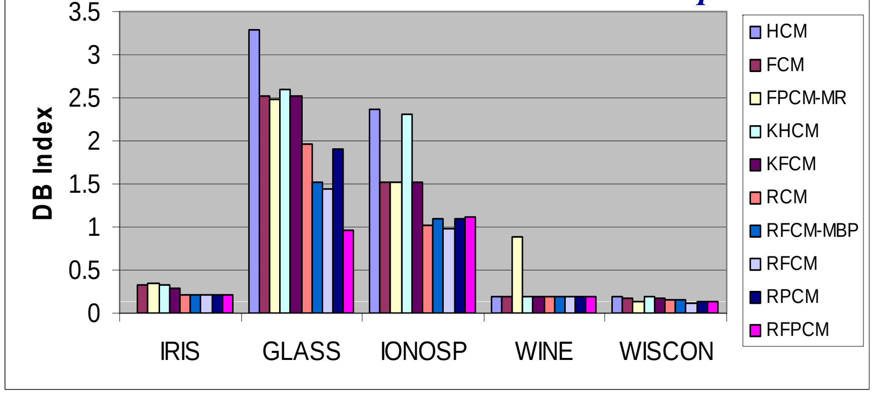
$$\beta = \frac{\mathrm{N}}{\mathrm{M}}; \quad \text{where} \quad \mathrm{N} = \sum_{i=1}^{c}\sum_{j=1}^{n_i} ||x_{ij} - \overline{v}||^2; \qquad \mathrm{M} = \sum_{i=1}^{c}\sum_{j=1}^{n_i} ||x_{ij} - v_i||^2; \quad \text{and} \quad \sum_{i=1}^{c} n_i = n;$$

For a given image and *c* value, ***higher the homogeneity within the segmented regions, higher would be β value***. The value also increases with *c*.

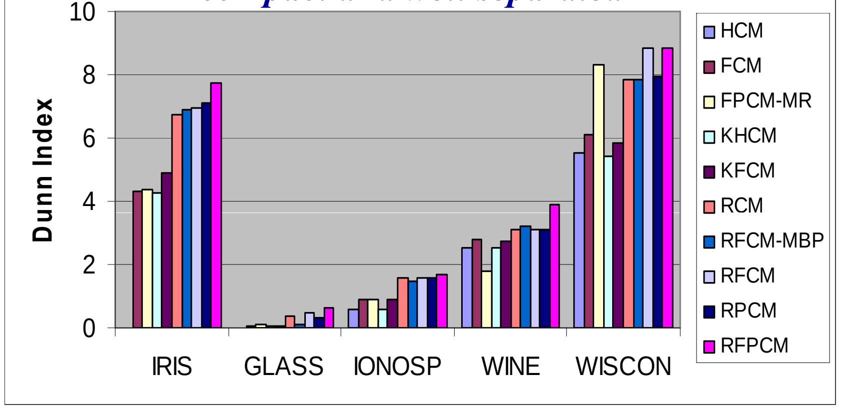
Clustering on Benchmark Data

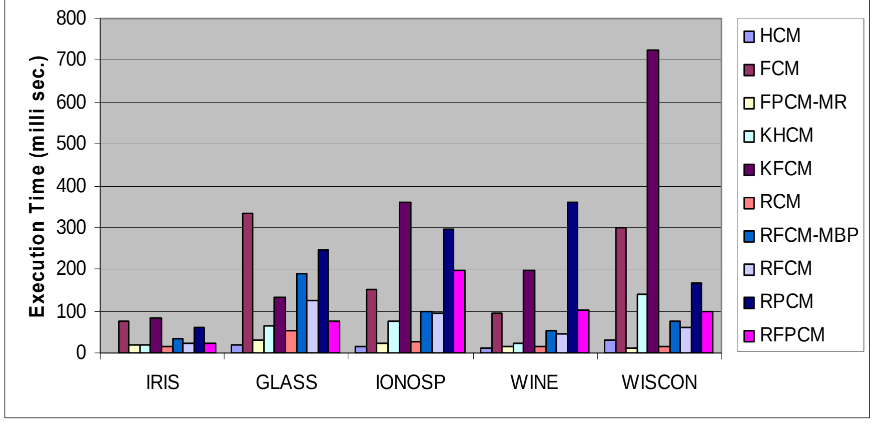
Data	Object	Feature	Cluster
Iris	150	4	3
Glass	214	10	6
Ionosp	351	34	2
Wine	178	13	3
Wiscon	569	30	2

within-cluster distance to between-cluster separation



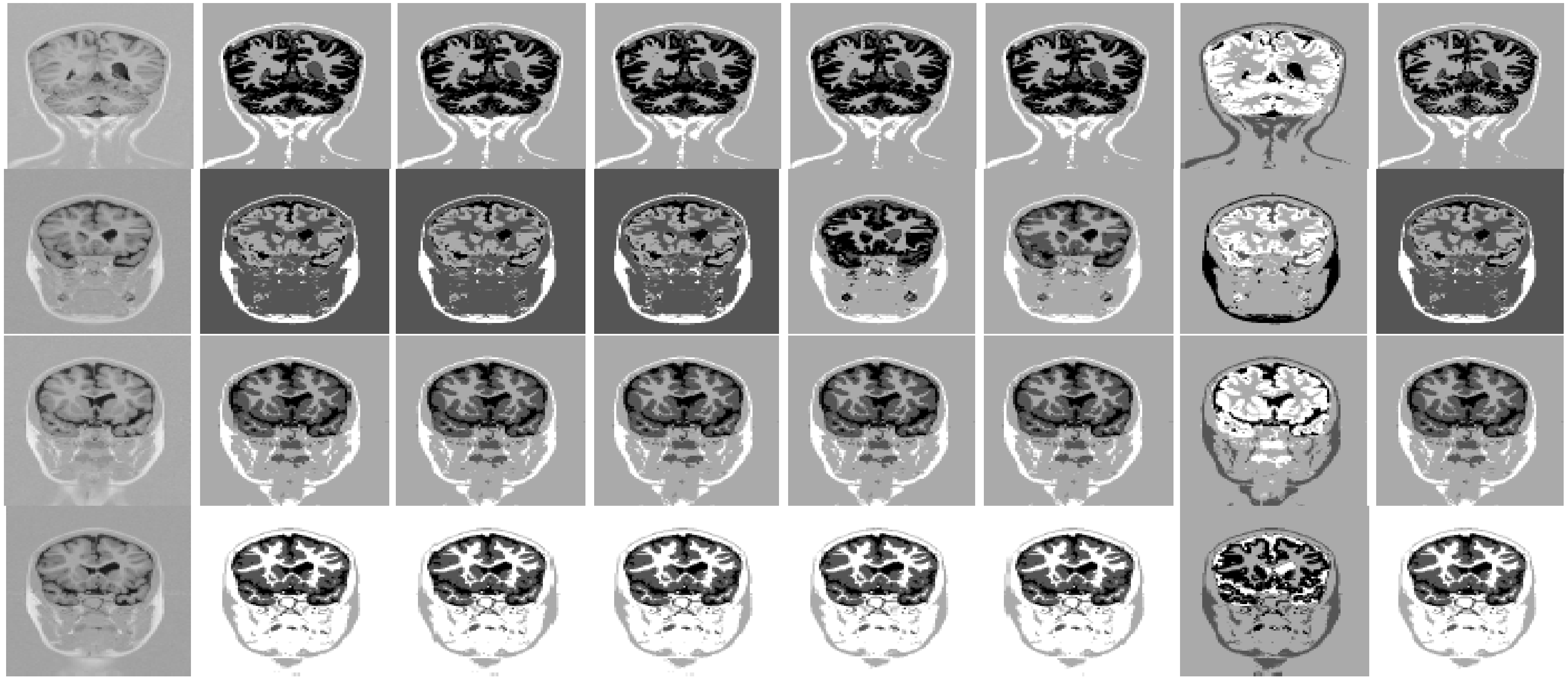
compact and well separated





RFPCM provides higher Dunn, and lower DB and execution time.

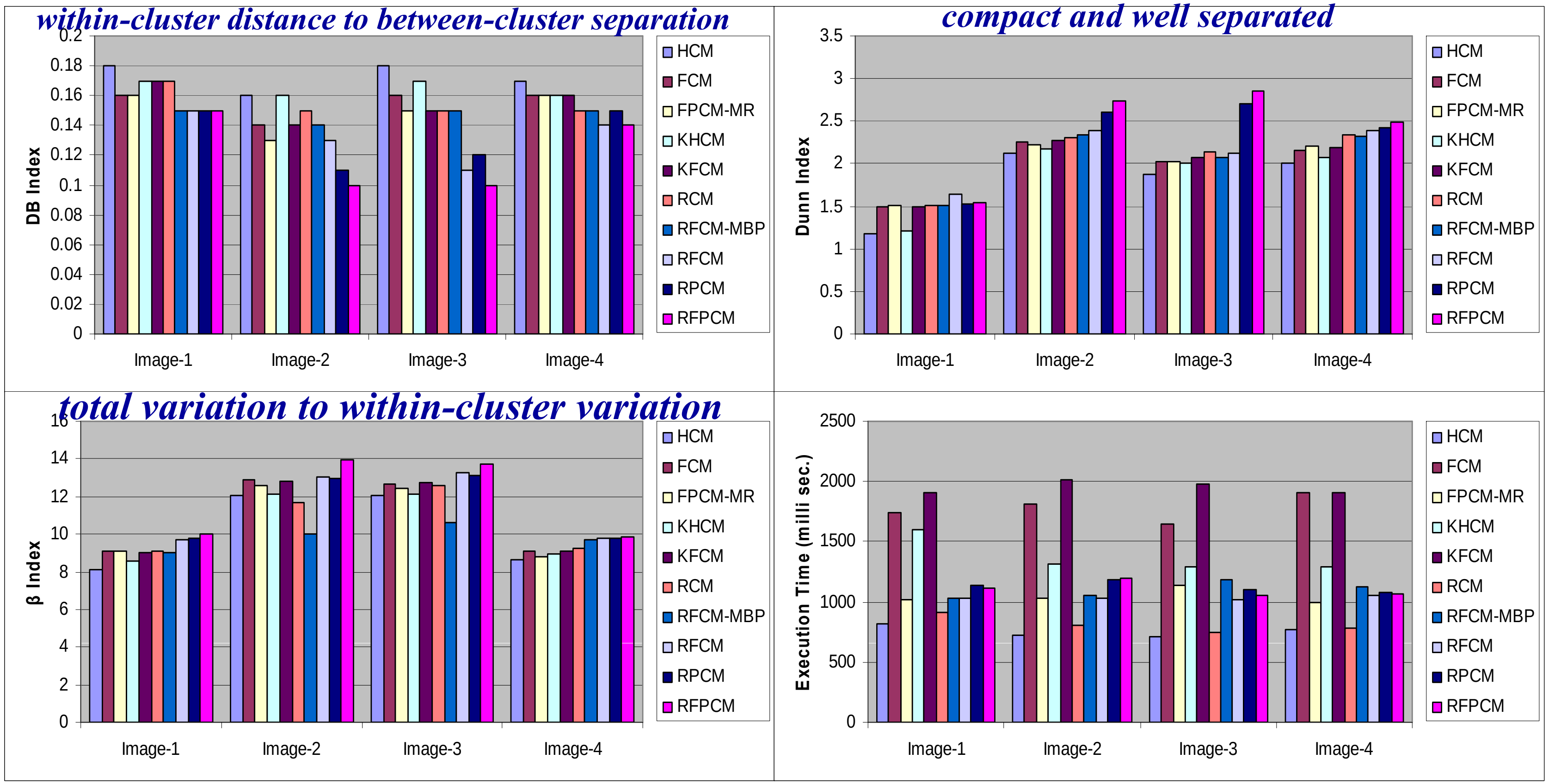
Brain MRI Segmentation (AMRI, Kolkata)



(a) Original (b) FCM (c) KFCM (d) RCM (e) RFCM^{MBP} (f) RFCM (g) RPCM (h) RFPCM

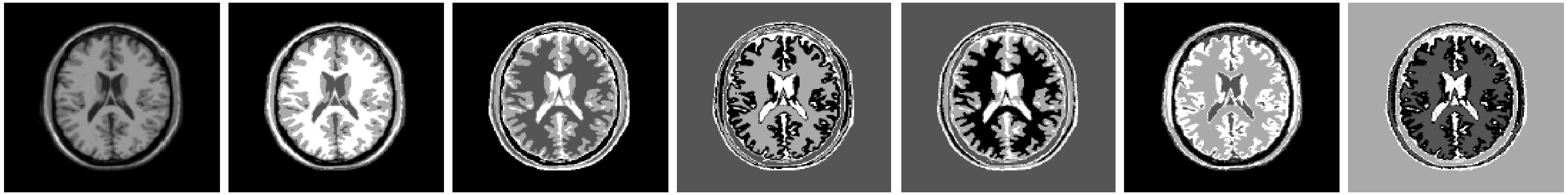
Each image is of size 256 X 180 with 16 bit gray levels: segmented into four categories, namely, CSF, gray matter, white matter, and background.

Brain MRI Segmentation (AMRI, Kolkata)



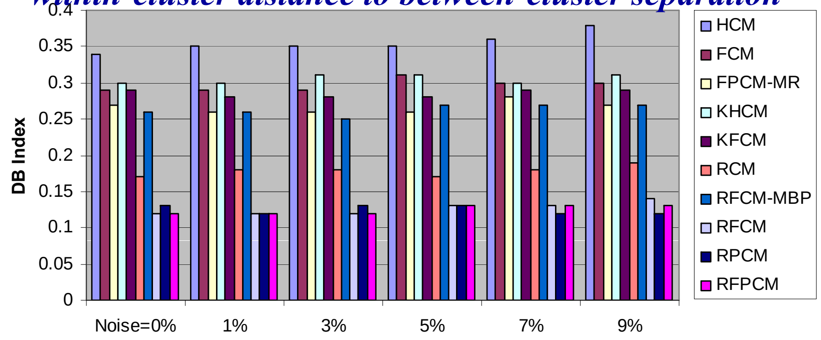
RFPCM provides lower DB, higher Dunn and β index, and lower execution time.

Brain MRI Segmentation (BrainWeb:SBD)

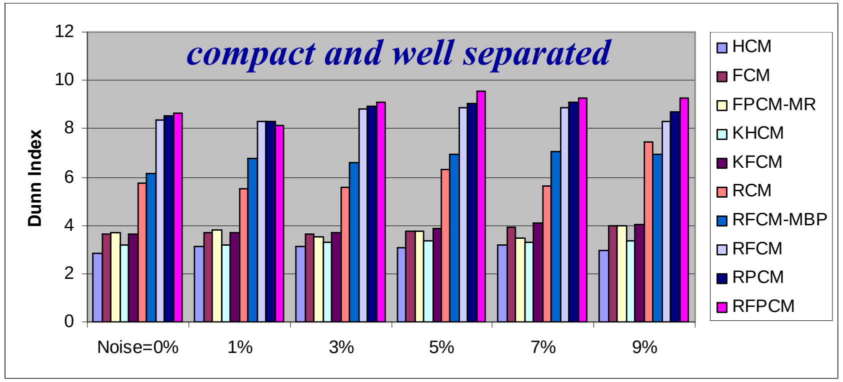


(a) Original (b) Noise=0% (c) Noise=1% (d) Noise=3% (e) Noise=5% (f) Noise=7% (g) Noise=9%

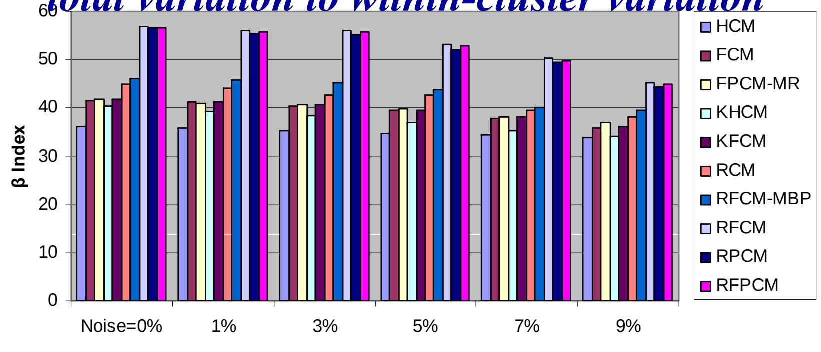
within-cluster distance to between-cluster separation



compact and well separated



total variation to within-cluster variation



The noise is calculated relative to the brightest tissue.

The performance of RFPCM is better than any other c-means algorithms.

Text-Graphics Segmentation Using HCM

Text-Graphics Segmentation Using FCM

Text-Graphics Segmentation Using PCM

Text-Graphics Segmentation Using FPCM

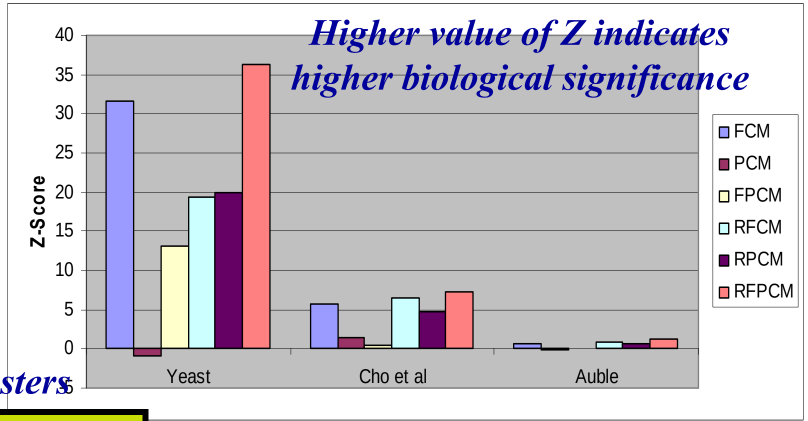
Text-Graphics Segmentation Using RFCM

Text-Graphics Segmentation Using RPCM

Text-Graphics Segmentation Using RFPCM

Clustering Functionally Similar Genes

Data Set	Genes	Condition	Cluster
Yeast	474	7	4
Cho et al	6457	17	33
Auble	6225	4	10



A low P-value indicates biologically significant gene clusters

FCM	RFPCM	Biological Function
2.81e^{-41}	2.39e^{-42}	Sporulation resulting in formation of cellular spore
7.24e^{-24}	1.02e^{-24}	Meiosis I
1.84e^{-16}	5.21e^{-21}	Ribosome Biogenesis
7.19e^{-10}	1.66e^{-11}	Glycosis

Highest Z-scores of 36.3 for yeast, 1.3 for Auble and 7.34 for Cho et al are obtained using RFPCM

Lowest P-value is achieved for RFPCM in case of yeast sporulation data

RFPCM provides more biologically significant clusters

Rough-Fuzzy C-Medoids Algorithm

- ***Cluster Prototypes (Medoids):*** $v_i = x_q$; where

$$q = \arg\max \begin{cases} w \times \mathcal{A} + \tilde{w} \times \mathcal{B} & \text{if } \underline{A}(\beta_i) \neq \emptyset, B(\beta_i) \neq \emptyset \\ \mathcal{A} & \text{if } \underline{A}(\beta_i) \neq \emptyset, B(\beta_i) = \emptyset \\ \mathcal{B} & \text{if } \underline{A}(\beta_i) = \emptyset, B(\beta_i) \neq \emptyset \end{cases}$$

$$\mathcal{A} = \sum_{x_k \in \underline{A}(\beta_i)} h(x_k, x_j);$$

$$\mathcal{B} = \sum_{x_k \in B(\beta_i)} (\mu_{ik})^{\acute{m}} h(x_k, x_j)$$

$$\mu_{ij} = \sum_{l=1}^{c} \left\{ \frac{h(x_j, v_i)}{h(x_j, v_l)} \right\}^{\frac{1}{\acute{m}-1}}$$

- Average normalized homology alignment scores of input subsequences with respect to their corresponding prototypes - *increases with increase in homology alignment scores within a cluster*

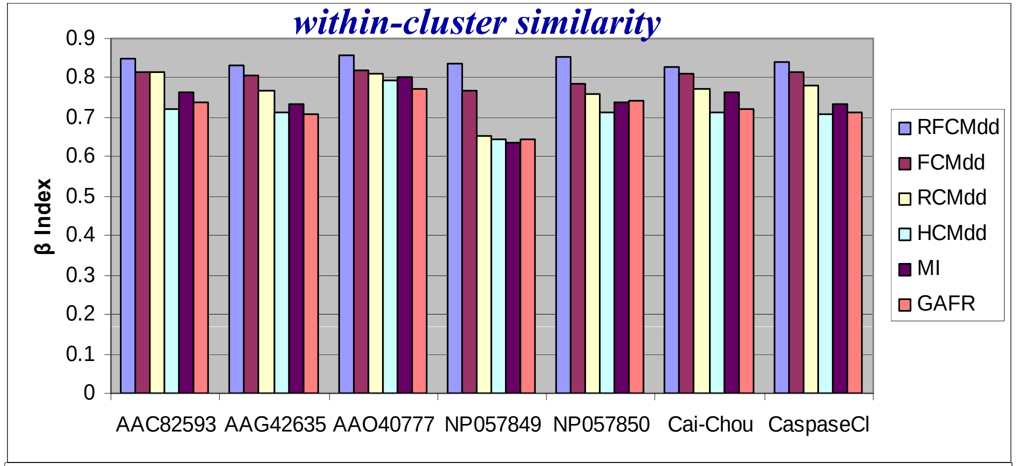
$$\beta = \frac{1}{c} \sum_{i=1}^{c} \frac{1}{n_i} \sum_{x_j \in \beta_i} \frac{h(x_j, v_i)}{h(v_i, v_i)}$$

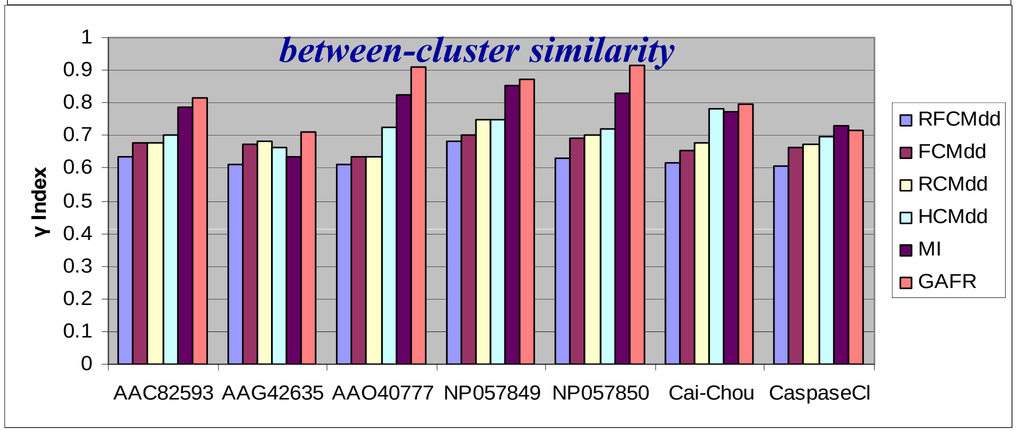
- Maximum normalized homology alignment score between prototypes - homology alignment score between all prototypes should be *as low as possible*

$$\gamma = \max_{i,j} \frac{1}{2} \left\{ \frac{h(v_j, v_i)}{h(v_i, v_i)} + \frac{h(v_i, v_j)}{h(v_j, v_j)} \right\}$$

P. Maji and S. K. Pal, ``Rough-Fuzzy *C*-Medoids Algorithm and Selection of Bio-Basis for Amino Acid Sequence Analysis'', *IEEE Trans. Knowledge and Data Engineering*, 19(6), pp. 859--872, 2007.

Amino Acid Sequence Clustering (NCBI)





- Rough-fuzzy c-medoids (RFCMdd): *pair-wise or relation clustering* - clusters are represented by medoids instead of means
- Dayhoff amino acid mutation matrix is used to compute similarity score
- ***RFCMdd provides higher β index and lower γ index values***

P. Maji and S. K. Pal, ``Rough-Fuzzy *C*-Medoids Algorithm and Selection of Bio-Basis for Amino Acid Sequence Analysis'', *IEEE Trans. Knowledge and Data Engineering*, 19(6), pp. 859--872, 2007.

Supervised Attribute Clustering

- Supervised attribute clustering: grouping of attributes, controlled by values of attributes as well as supervised information of sample categories

- To find groups of co-regulated genes whose collective expression is strongly associated with sample categories

$$\boxed{\psi(\mathcal{A}_i, \mathcal{A}_j) = 1 - \kappa}$$

Fuzzy-rough supervised similarity measure between two condition attributes A_i and A_j

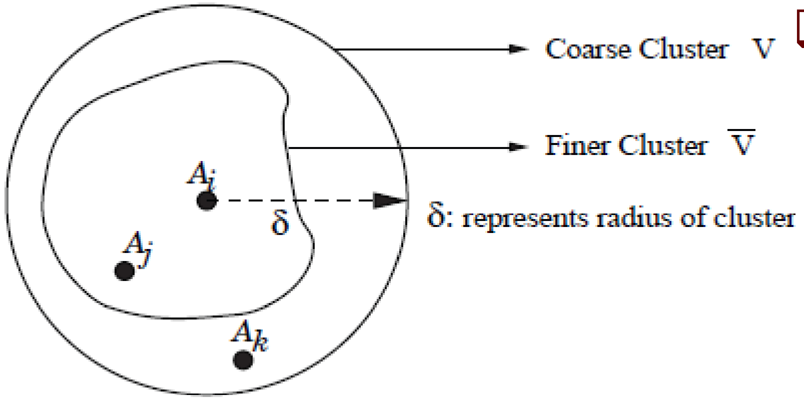
$$\text{where} \quad \kappa = \left\{ \frac{\sigma_{\{\mathcal{A}_i,\mathcal{A}_j\}}(\mathbb{D}, \mathcal{A}_j) + \sigma_{\{\mathcal{A}_i,\mathcal{A}_j\}}(\mathbb{D}, \mathcal{A}_i)}{2} \right\} = \gamma_{\{\mathcal{A}_i,\mathcal{A}_j\}}(\mathbb{D}) - \left\{ \frac{\gamma_{\mathcal{A}_i}(\mathbb{D}) + \gamma_{\mathcal{A}_j}(\mathbb{D})}{2} \right\}$$

$$\sigma_{\{\mathcal{A}_i,\mathcal{A}_j\}}(\mathbb{D}, \mathcal{A}_j) = \gamma_{\{\mathcal{A}_i,\mathcal{A}_j\}}(\mathbb{D}) - \gamma_{\mathcal{A}_i}(\mathbb{D}).$$

Significance of fuzzy condition attributes A_j with respect to A_i

Fuzzy-Rough Supervised Clustering

- Proposed supervised attribute clustering:
  - determining the relevance of each attribute;
  - growing the cluster around each relevant attribute incrementally by adding one attribute after the other



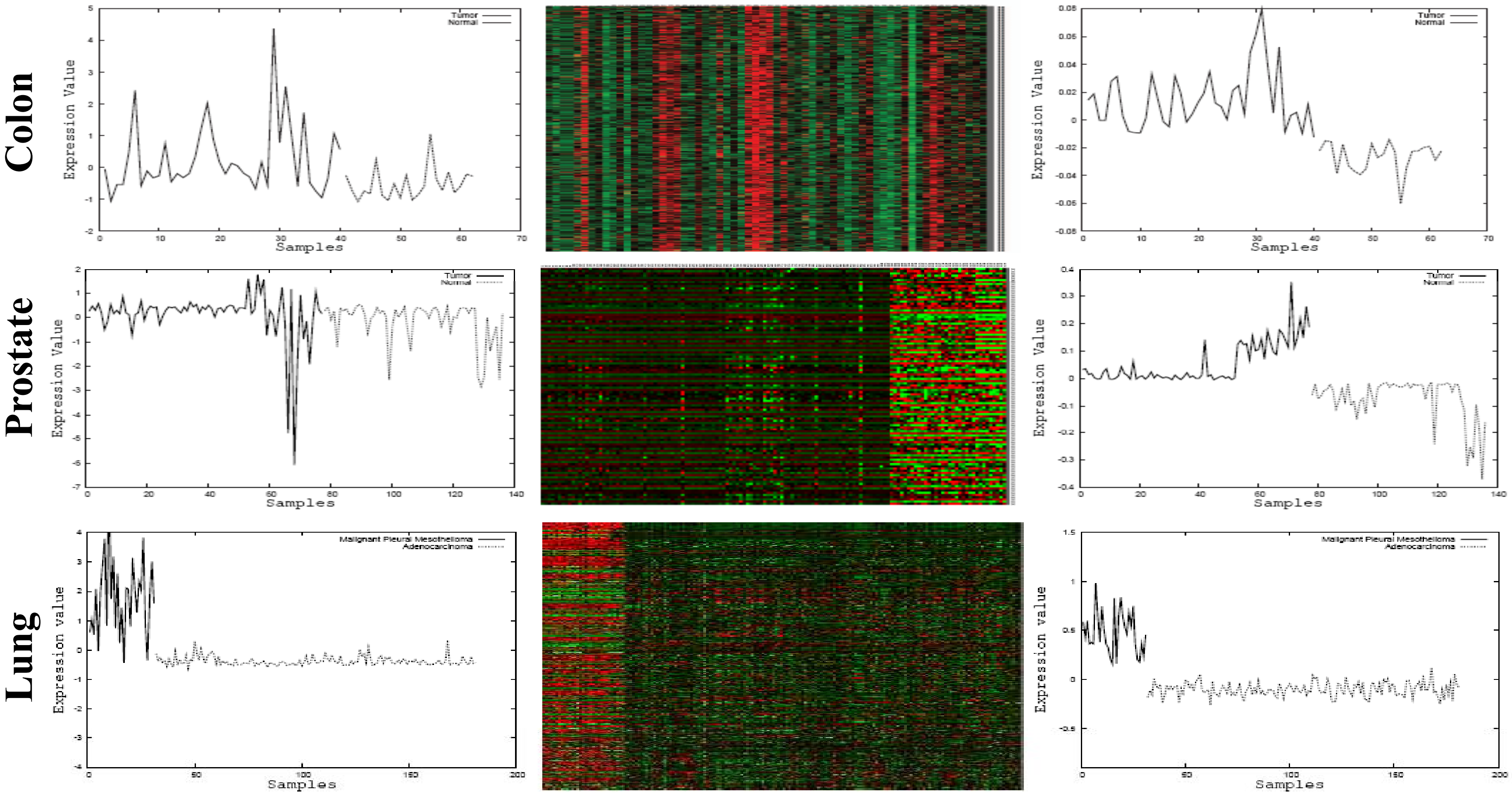
- Each cluster is augmented by attributes that satisfy following two conditions:
  - suit best into the current cluster in terms of a supervised similarity measure (*coarse*);
  - improve differential expression of current cluster most, according to the relevance of cluster representative (*finer*).

$$\mathbb{V}_i = \{\mathcal{A}_j | \psi(\mathcal{A}_i, \mathcal{A}_j) \geq \delta \; \mathcal{A}_j \neq \mathcal{A}_i \in \mathbb{C}\}.$$

$\bar{\mathbb{V}}_i \subset \mathbb{V}_i$ - increase the differential expression of cluster representative

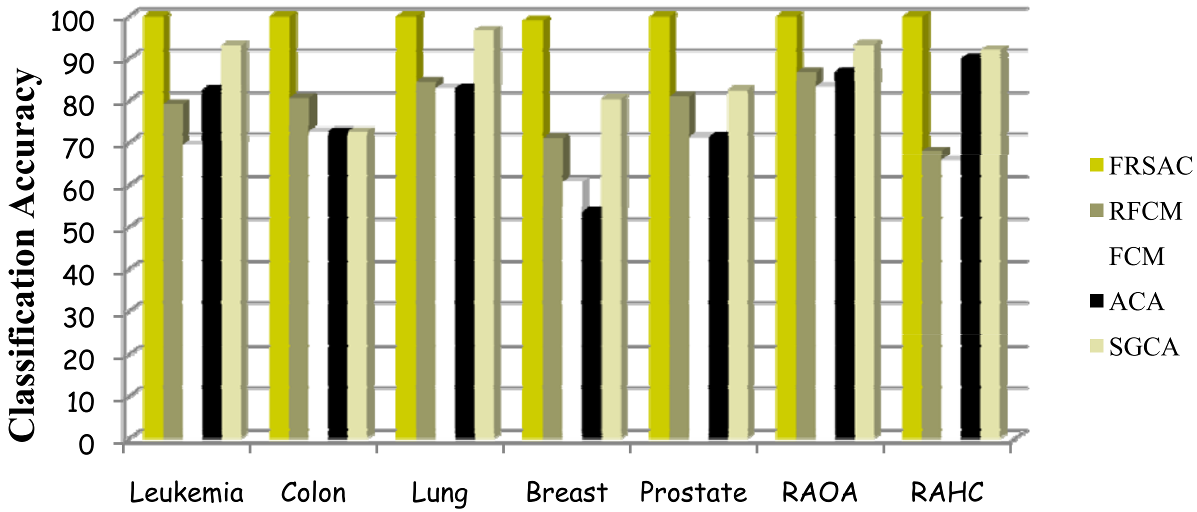
P. Maji, ``Fuzzy-Rough Supervised Attribute Clustering Algorithm and Classification of Microarray Data'', *IEEE Trans. System, Man and Cybernetics, Part B, Cybernetics*, 41(1), pp. 222--233, 2011.

Co-regulated Clusters of Genes



(a) Initial expression value (b) Eisen plot (c) Augmented expression value

Performance Analysis



FRSAC provides higher classification accuracy using support vector machine

Biological Significance Analysis

Data Sets	Methods / Algorithms	Biological Process		
		Gene Ontology Term	p-value	FDR
Breast	Proposed	Positive regulation of biological process	5.7E-027	0
	GS	*	*	*
	SGCA	*	*	*
Leukemia	Proposed	Multicellular organismal development	2.1E-02	10
	GS	*	*	*
	SGCA	*	*	*
Colon	Proposed	Cellular process	1.8E-012	0
	GS	Regulation of system process	1.2E-02	32
	SGCA	Blood vessel development	1.7E-02	20
RAOA	Proposed	Immune system process	8.3E-07	0
	GS	Immune system process	2.9E-016	0
	SGCA	Interspecies interaction between organisms	7.4E-03	18

Conclusion and Discussion

- Rough-fuzzy clustering - combination of *restrictive* (hard) and *descriptive* (fuzzy) clustering
  - Generalization of existing clustering algorithms
  - Rough-fuzzy *c*-medoids (relational or pair-wise clustering) – applicable for relational data analysis of bioinformatics, data and web mining
  - Rough-fuzzy supervised attribute clustering - co-regulated gene clusters

- Rough-fuzzy feature selection – relevance, significance, and redundancy

- The emphasis of these methodologies is given on
  - handling large data sets, both in size and dimension, and involve classes that are overlapping, intractable and/or having nonlinear boundaries;
  - demonstrating their success in certain tasks of bioinformatics and medical imaging.

References

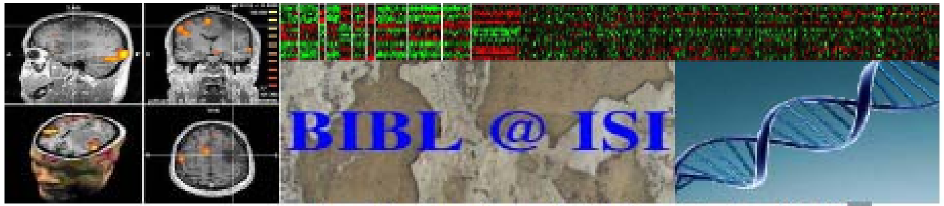
- J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithm*. New York: Plenum, 1981.
- R. Krishnapuram and J. M. Keller, "A Possibilistic Approach to Clustering," *IEEE Trans. Fuzzy Systems*, vol. 1, no. 2, pp. 98–110, 1993.
- N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, "A Possibilistic Fuzzy C-Means Clustering Algorithm," *IEEE Trans. Fuzzy Systems*, vol. 13, no. 4, pp. 517–530, 2005.
- F. Masulli and S. Rovetta, "Soft Transition From Probabilistic to Possibilistic Fuzzy Clustering," *IEEE Trans. Fuzzy Systems*, vol. 14, no. 4, pp. 516–527, 2006.
- P. Lingras and C. West, "Interval Set Clustering of Web Users with Rough K-means," *Journal of Intelligent Information Systems*, vol. 23, no. 1, pp. 5–16, 2004.
- S. Mitra, H. Banka, and W. Pedrycz, "Rough-Fuzzy Collaborative Clustering," *IEEE Trans. Systems, Man, and Cybernetics - Part B*, vol. 36, pp. 795–805, 2006.
- J. C. Bezdek and N. R. Pal, "Some New Indexes for Cluster Validity," *IEEE Trans. System, Man, and Cybernetics, Part B*, vol. 28, pp. 301–315, 1988.
- S. K. Pal, A. Ghosh, and B. U. Sankar, "Segmentation of Remotely Sensed Images with Fuzzy Thresholding, and Quantitative Evaluation," *International Journal of Remote Sensing*, vol. 21, no. 11, pp. 2269–2300, 2000.
- Z. Pawlak, *Rough Sets, Theoretical Aspects of Reasoning About Data*. Dordrecht, The Netherlands: Kluwer, 1991.
- Z.R. Yang and R. Thomson, "Bio-Basis Function Neural Network for Prediction of Protease Cleavage Sites in Proteins," *IEEE Trans. Neural Networks*, vol. 16, no. 1, pp. 263-274, 2005.

Some Related Publications

- Pradipta Maji and Sankar K. Pal, "*Rough-Fuzzy Pattern Recognition: Applications in Bioinformatics and Medical Imaging*", John Wiley & Sons, Inc., New Jersey/IEEE Computer Society Press, ISBN: 978-1-1180-0440-1, January, 2012.
- Pradipta Maji and Sushmita Paul, "Rough-Fuzzy Clustering for Grouping Functionally Similar Genes from Microarray Data", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10(2), pp. 286--299, 2013.
- Pradipta Maji, "Fuzzy-Rough Supervised Attribute Clustering Algorithm and Classification of Microarray Data", *IEEE Transactions on System, Man and Cybernetics, Part B, Cybernetics*, 41(1), pp. 222--233, 2011.
- Pradipta Maji and Sankar K. Pal, "Rough Set Based Generalized Fuzzy *C*-Means Algorithm and Quantitative Indices'', *IEEE Transactions on System, Man and Cybernetics, Part B, Cybernetics*, 37(6), pp. 1529--1540, 2007.
- Pradipta Maji and Sankar K. Pal, "Rough-Fuzzy *C*-Medoids Algorithm and Selection of Bio-Basis for Amino Acid Sequence Analysis'', *IEEE Transactions on Knowledge and Data Engineering*, 19(6), pp. 859--872, 2007.

Biomedical Imaging and Bioinformatics Lab (BIBL)

Biomedical Imaging and Bioinformatics Lab http://www.isical.ac.in/~bibl/



Biomedical Imaging and Bioinformatics Lab (BIBL)

- BIBL
- BIBL Members
- Publications
- Research
- Projects
- Software
- Alumnis
- BIBL Info
- Contact

Welcome to BIBL

Biomedical Imaging and Bioinformatics Lab (BIBL) is formally established in the academic year 2012-2013 to carry out fundamental and advanced research in the fields of biomedical imaging and bioinformatics. Aim of the BIBL is to conduct basic research in pattern recognition, machine learning, soft computing, medical image analysis and processing, computational biology and bioinformatics.

The Lab has been set up at the Ground Floor of S. N. Bose Bhavan (Library Building), Indian Statistical Institute (ISI), Kolkata, India.

Contact

Biomedical Imaging and Bioinformatics Lab (BIBL)
Ground Floor, S. N. Bose Bhaban
Indian Statistical Institute (ISI)
203 B. T. Road, Kolkata, 700108, INDIA

Phone: +91 33 2575 XXXX
Fax: +91 33 XXXX XXXX

E-mail: bibl@isical.ac.in
URL: http://www.isical.ac.in/~bibl

Acknowledgment: BIBL Members

Thank you