Rough Sets and Decision Making

JOUNI JÄRVINEN
ALEKSANDR MYLLÄRI

Department of Information Technology
University of Turku
Finland

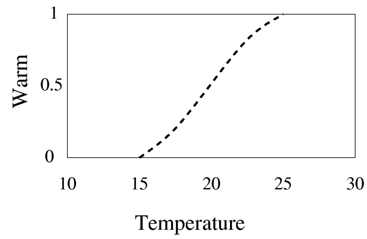
Petrozavodsk
June 26$^{\text{th}}$ 2003

INTRODUCTION

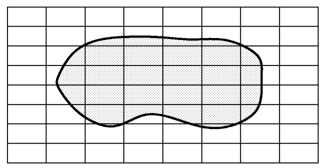
- Introduced by Z. Pawlak in the early eighties
- Deals with uncertainty in data

Rough sets ≠ Fuzzy sets

Fuzzy sets:
Fuzzy membership function $U \to [0, 1]$



Rough sets:
Approximations based on an indiscernibility relation



SOME APPLICATIONS PRESENTED IN RSCTC CONFERENCES (1998, 2000, 2002)

- EEG analysis
- Medical diagnostic rules
- Post-surgery survival analysis
- Similarity of DNA sequences
- cDNA microarray analysis
- Robot navigation
- Satellite attitude control
- Switchbox routing
- Job scheduling
- Industrial design
- Business process understanding
- Visual classification
- Text classification
- Web mining
- Identification of low-paying workplaces
- Monitoring loose parts in nuclear power plants

INDISCERNIBILITY RELATIONS

Let us denote

$$x \approx y$$

if we cannot discern x and y by their properties.

- $\approx$ is called an **indiscernibility relation** on U

Usually indiscernibility relations are assumed to be **equivalences:**

- $x \approx x$ (reflexive)
- $x \approx y \Rightarrow y \approx x$ (symmetric)
- $x \approx y$ and $y \approx z \Rightarrow x \approx z$ (transitive)

The **equivalence class** of $x \in U$

$$[x] = \{y \in U \mid x \approx y\}$$

consists of objects indiscernible from x

The set $U/\!\approx$ of all equivalence classes forms a **partition** of U.

INDISCERNIBILITY AND DECISION TABLES

A **decision table** is a triple $\mathcal{T} = (U, C, D)$

- U is a set of **objects**
- C is a set of **condition attributes**
- D is a set of **decision attributes**
- each attribute $a \in C \cup D$ is a map $a: U \rightarrow V_a$
- V_a is the **value set** of the attribute a

For any $B \subseteq C$, we may define a binary relation

$$\mathrm{ind}(B) = \{(x, y) \in U \times U \mid (\forall a \in B)\ a(x) = a(y)\}$$

- $\mathrm{ind}(B)$ is called the B-**indiscernibility** relation

If $(x, y) \in \mathrm{ind}(B)$, then objects x and y are indiscernible with respect to all attributes in B

EXAMPLE OF INDISCERNIBILITY

U	HEADACHE	TEMPERATURE	FLU
1	yes	normal	no
2	yes	high	yes
3	yes	normal	no
4	yes	very high	no
5	no	high	no
6	no	very high	yes
7	no	high	yes
8	no	very high	yes

The partition by HEADACHE:

$$\{\{1, 2, 3, 4\}, \{5, 6, 7, 8\}\}$$

The partition by TEMPERATURE:

$$\{\{1, 3\}, \{2, 5, 7\}, \{4, 6, 8\}\}$$

The partition by HEADACHE and TEMPERATURE:

$$\{\{1, 3\}, \{2\}, \{4\}, \{5, 7\}, \{6, 8\}\}$$

ROUGH APPROXIMATIONS

Let $\approx$ be an indiscernibility relation on U

Lower approximation of $X \subseteq U$:

$$X^{\blacktriangledown} = \{x \in U \mid\ [x] \subseteq X\}$$

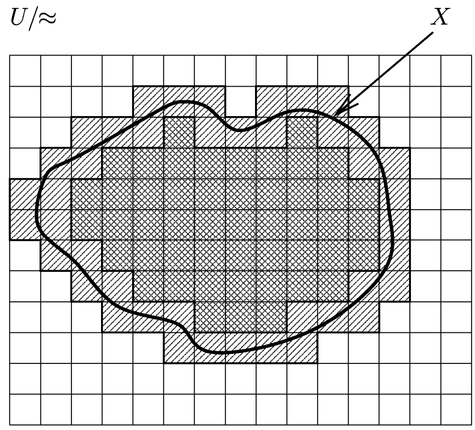
Upper approximation of $X \subseteq U$:

$$X^{\blacktriangle} = \{x \in U \mid\ X \cap [x] \neq \emptyset\}$$

Boundary of $X \subseteq U$:

$$B(X) = X^{\blacktriangle} - X^{\blacktriangledown}$$

ROUGH APPROXIMATIONS



Lower approximation: 
- Elements which certainly are in X

Upper approximation:  + 
- Elements which possibly are in X

Boundary: 
- Area of uncertainty

Definability and Accuracy of Approximation

X is **definable** $\overset{\text{def}}{\iff} X^{\blacktriangle} = X^{\blacktriangledown} \iff B(X) = \emptyset$

The definable sets are $\emptyset$ and the unions of equivalence classes of $\approx$.

Accuracy of approximation:

$$\alpha(X) = \frac{|X^{\blacktriangledown}|}{|X^{\blacktriangle}|} \qquad (X \neq \emptyset)$$

- If X is definable, then $\alpha(X) = 1$
- In particular, $\alpha(X) = 0 \iff X^{\blacktriangledown} = \emptyset$

Example of Approximations

Let

$$X = \{x \in U \mid \text{FLU}(x) = \text{yes}\} = \{2, 6, 7, 8\}$$

The partition by HEADACHE and TEMPERATURE:

$$\{\{1, 3\}, \{2\}, \{4\}, \{5, 7\}, \{6, 8\}\}$$

Approximations:

$X^{\blacktriangledown}$	$=$	$\{2, 6, 8\}$	— certainly
$X^{\blacktriangle}$	$=$	$\{2, 5, 6, 7, 8\}$	— possibly
$B(X)$	$=$	$\{5, 7\}$	— area of uncertainty
$(X^{\blacktriangle})^{\complement}$	$=$	$\{1, 3, 4\}$	— certainly not

Accuracy of approximation:

$$\alpha(X) = \frac{|X^{\blacktriangledown}|}{|X^{\blacktriangle}|} = \frac{3}{5} = 0.6$$

Basic Properties of Approximations

(a) Approximations are definable

(b) $X^{\blacktriangledown} \subseteq X \subseteq X^{\blacktriangle}$

(c) $\emptyset^{\blacktriangledown} = \emptyset^{\blacktriangle} = \emptyset$ and $U^{\blacktriangledown} = U^{\blacktriangle} = U$

(d) $X^{\blacktriangle} \cup Y^{\blacktriangle} = (X \cup Y)^{\blacktriangle}$

(e) $X^{\blacktriangledown} \cap Y^{\blacktriangledown} = (X \cap Y)^{\blacktriangledown}$

(f) $(X^{\complement})^{\blacktriangle} = (X^{\blacktriangledown})^{\complement}$ and $(X^{\complement})^{\blacktriangledown} = (X^{\blacktriangle})^{\complement}$

(g) If $X \subseteq Y$, then $X^{\blacktriangledown} \subseteq Y^{\blacktriangledown}$ and $X^{\blacktriangle} \subseteq Y^{\blacktriangle}$

(h) Definable sets form a complete field of sets

Rough Sets

Rough equality relation:

$$X \equiv Y \overset{\text{def}}{\iff} X^{\blacktriangledown} = Y^{\blacktriangledown} \text{ and } X^{\blacktriangle} = Y^{\blacktriangle}$$

The relation $\equiv$ is an equivalence on $\wp(U)$

The equivalence classes of $\equiv$ are called **rough sets**

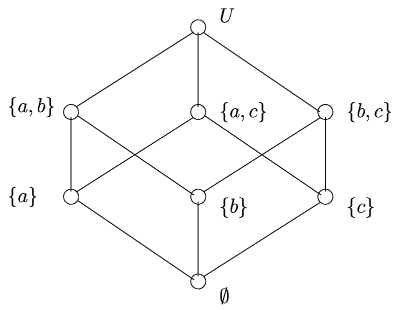
Idea: If different subsets of U are observed through the knowledge represented by the indiscernibility relation $\approx$, the sets in the same rough set look the same

STRUCTURE OF SETS

The set $\wp(U)$ of all subsets of U ordered with $\subseteq$ is isomorphic to $\mathbf{2}^U$, where

2

Example. Let $U = \{a, b, c\}$



$\cong \mathbf{2}^3$

STRUCTURE OF ROUGH SETS

Each rough set can be viewed as a pair $(X^\blacktriangledown, X^\blacktriangle)$

The set $\mathcal{R}$ of rough sets can be ordered by

$$(X^\blacktriangledown, X^\blacktriangle) \leq (Y^\blacktriangledown, Y^\blacktriangle) \overset{\text{def}}{\iff} X^\blacktriangledown \subseteq Y^\blacktriangledown \text{ and } X^\blacktriangle \subseteq Y^\blacktriangle$$

$(\mathcal{R}, \leq)$ is a Stone algebra:

- bounded distributive lattice
- each element x has a pseudocomplement x^*

$$\boxed{x \wedge x^* = 0 \text{ and } x \wedge a = 0 \Rightarrow a \leq x^*}$$

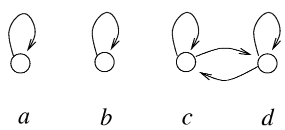
Moreover, $(\mathcal{R}, \leq)$ is isomorphic to

$$\mathbf{2}^I \times \mathbf{3}^J$$

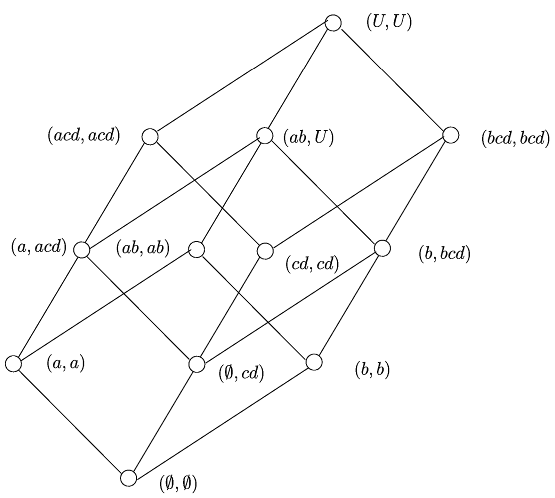
where

- $I = \{\ [x]\ \mid \text{cardinality of } [x] = 1\}$
- $J = \{\ [x]\ \mid \text{cardinality of } [x] > 1\}$

EXAMPLE. Let $U = \{a, b, c, d\}$



The lattice $(\mathcal{R}, \leq)$ of all rough sets:

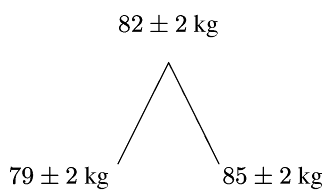


$\cong \mathbf{2}^2 \times \mathbf{3}$

ROUGH SETS DEFINED BY TOLERANCES

There are *non*-transitive indiscernibility relations

Example. Weight of three persons:



J. JÄRVINEN, *Approximations and Rough Sets Based on Tolerances*, in: W. ZIARKO, Y. YAO (eds.), Proceedings of RSCTC 2000, LNAI **2005** (Springer-Verlag, Berlin 2001) pp. 182–189.

Generalizations of Approximations

Let $\mathcal{B} = (B, \leq)$ be a complete atomic Boolean lattice.
A map $\varphi\colon \mathcal{A}(\mathcal{B}) \to B$ is

extensive: $x \leq \varphi(x)$
symmetric: $x \leq \varphi(y) \;\Rightarrow\; y \leq \varphi(x)$
closed: $x \leq \varphi(y) \;\Rightarrow\; \varphi(x) \leq \varphi(y)$

Generalized approximations:

$$x^{\blacktriangledown} = \bigvee \{a \in \mathcal{A}(\mathcal{B}) \mid \varphi(a) \leq x\}$$
$$x^{\blacktriangle} = \bigvee \{a \in \mathcal{A}(\mathcal{B}) \mid \varphi(a) \wedge x \neq 0\}$$

J. JÄRVINEN, *On the Structure of Rough Approximations*, Fundamenta Informaticae **53** (2002) pp. 135–153.

Different Information Relations

Each attribute is a map $a\colon U \to \wp(V_a)$

Strong similarity:

$$(\forall a \in A)\; a(x) \cap a(y) \neq \emptyset$$

Weak inclusion:

$$(\exists a \in A)\; a(x) \subseteq a(y)$$

Strong negative similarity:

$$(\forall a \in A)\; a(x)^{\mathsf{C}} \cap a(y)^{\mathsf{C}} \neq \emptyset$$

Weak incomplementarity:

$$(\exists a \in A)\; a(x) \neq a(y)^{\mathsf{C}}$$

J. JÄRVINEN, *Preimage Relations and Their Matrices*. In L. POLKOWSKI, A. SKOWRON (eds.), Proceedings of RSCTC 1998, LNAI **1424**, (Springer-Verlag, Berlin 1998), pp. 139–146.

Dependency Relations

It may be possible that the values of some attribute set are determined by another set of attributes:

$$X \to Y \overset{\text{def}}{\iff} \text{ind}(X) \subseteq \text{ind}(Y)$$

- J. JÄRVINEN, *A Representation of Dependence Spaces and Some Basic Algorithms*, Fundamenta Informaticae **29** (1997), pp. 369–382.
- J. JÄRVINEN, *Difference Functions of Dependence Spaces*, Acta Cybernetica **14** (2000) pp. 619–630.
- J. JÄRVINEN, *Armstrong Systems on Ordered Sets*, in: C.S. CALUDE, M.J. DINNEEN, S. SBURLAN (eds.), Combinatorics, Computability, Logic (Springer-Verlag, London 2001) pp. 137–149.