Rough Sets: Theoretical aspects of reasoning about data

wu

July 10, 2019

Contents

1 Knowledge

1.1 Knowledge base

Given a finite set $U \neq \emptyset$ (the universe). Any subset $X \subset U$ of the universe is called a **concept** or a **category** in U. And any family of concepts in U will be referred to as **abstract knowledge** about U.

partition or **classification** of a certain universe U is a family $C = \{X_1, X_2, \dots, X_n\}$ s.t. $X_i \subset U, X_i \neq \emptyset, X_i \cap X_j = \emptyset$ and $\bigcup X_i = U$

A family of classifications is called a **knowledge base** over U

R an equivalence relation over U, U/R family of all equivalence classes of R, referred to be **categories** or **concepts** of R, and $[x]_R$ denotes a category in R containing an element $x \in U$

By a **knowledge base** we can understand a relational system $K = (U, \boldsymbol{R})$, $\boldsymbol{R}$ is a family of equivalence relations over U

If $\boldsymbol{P} \subset \boldsymbol{R}$ and $\boldsymbol{P} \neq \emptyset$, then $\bigcap \boldsymbol{P}$ is also an equivalence relation, and will be denoted by $IND(\boldsymbol{P})$, called an **indiscernibility relation** over $\boldsymbol{P}$

$$[x]_{IND(\boldsymbol{P})} = \bigcap_{R \in \boldsymbol{P}} [x]_R$$

$U/IND(\boldsymbol{P})$ called **$\boldsymbol{P}$-basic knowledge about** U in K. For simplicity, $U/\boldsymbol{P} = U/IND(\boldsymbol{P})$ and $\boldsymbol{P}$ will be also called **$\boldsymbol{P}$-basic knowledge** . Equivalence classes of $IND(\boldsymbol{P})$ are called **basic categories** of knowledge $\boldsymbol{P}$. If $Q \in \boldsymbol{R}$, then Q is a **Q-elementary knowledge** and equivalence classes of Q are referred to as **Q-elementary concepts** of knowledge $\boldsymbol{R}$

The family of all $\boldsymbol{P}$-basic categories for all $\neq \boldsymbol{P} \subset \boldsymbol{R}$ will be called the **family of basic categories** in knowledge base $K = (U, \boldsymbol{R})$

Let $K = (U, \boldsymbol{R})$ be a knowledge base. By $IND(K)$ we denote the family of all equivalence relations defined in K as $IND(K) = \{IND(\boldsymbol{P}) : \emptyset \neq \boldsymbol{P} \subseteq \boldsymbol{R}\}$.

Thus $IND(K)$ is the minimal set of equivalence relations.

Every union of $\boldsymbol{P}$-basic categories will be $\boldsymbol{P}$**-category**

The family of all categories in the knowledge base $K = (U, \boldsymbol{R})$ will be referred to as K**-categories**

1.2 Equivalence, generalization and specialization of knowledge

Let $K = (U, \boldsymbol{P}), K' = (U, \boldsymbol{Q})$. K and K' are **equivalent** $K \simeq K', (\boldsymbol{P} \simeq \boldsymbol{Q})$ if $IND(\boldsymbol{P}) = IND(\boldsymbol{Q})$. Hence $K \simeq K'$ if both K and K' have the same set of elememtary categories. *This means that knowledge in knowledge bases K and K' enables us to express exactly the same facts about the universe.*

If $IND(\boldsymbol{P}) \subset IND(\boldsymbol{Q})$ then knowledge $\boldsymbol{P}$ is **finer** than knowledge $\boldsymbol{Q}$ (**coarser**). $\boldsymbol{P}$ is **specialization** of $\boldsymbol{Q}$ and $\boldsymbol{Q}$ is **generalization** of $\boldsymbol{P}$

2 Imprecise categories, approximations and rough sets

2.1 Rough sets

Let $X \subseteq U$. X is R**-definable** or R**-exact** if X is the union of some R-basic categories. otherwise R**-undefinable**, R**-rough**, R**-inexact** .

2.2 Approximations of set

Given $K = (U, \boldsymbol{R}), R \in IND(K)$

$$\underline{R}X = \bigcup \{Y \in U/R : Y \subseteq X\}$$
$$\overline{R}X = \bigcup \{Y \in U/R : Y \cap X \neq \emptyset\}$$

called the R**-lower** and R**-upper approximation** of X

$BN_R(X) = \overline{R}X - \underline{R}X$ is R**-boundary** of X. $BN_R(X)$ is the set of elements which cannot be classified either to X or to $-X$ having knowledge R

$$POS_R(X) = \underline{R}X, R\text{-positive region of } X$$
$$NEG_R(X) = U - \overline{R}X, R\text{-negative region of } X$$
$$BN_R(X) - R\text{-borderline region of } X$$

If $x \in POS(X)$, then x will be called an **R-positive example of X**

Proposition 2.1. *1. X is R-definable if and only if $\underline{R}X = \overline{R}X$*

2. X is rought w.r.t. R if and only if $\underline{R}X \neq \overline{R}X$

2.3 Properties of approximations

Proposition 2.2 (2.2)**.** *1. $\underline{R}X \subseteq X \subseteq \overline{R}X$*

2. $\underline{R}\emptyset = \underline{R}\emptyset = \emptyset; \quad \underline{R}U = \overline{R}U = U$

3. $\overline{R}(X \cup Y) = \overline{R}X \cup \overline{R}Y$

4. $\underline{R}(X \cap Y) = \underline{R}X \cap \underline{R}Y$

5. $X \subseteq Y$ implies $\underline{R}X \subseteq \underline{R}Y$

6. $X \subseteq Y$ implies $\overline{R}X \subseteq \overline{R}Y$

7. $\underline{R}(X \cup Y) \subseteq \underline{R}X \cup \underline{R}Y$

8. $\underline{R}(-X) = -\overline{R}X$

9. $\overline{R}(-X) = -\underline{R}X$

10. $\overline{R}(-X) = -\underline{R}X$

11. $\underline{RR}X = \overline{R}\underline{R}X = \underline{R}X$

12. $\overline{RR}X = \underline{R}\overline{R}X = \overline{R}X$

The equivalence relation R over U uniquely defines a topological space $T = (U, DIS(R))$ where $DIS(R)$ is the familty of all open and closed set in T and U/R is a base for T. The R-lower and R-upper approximation of X in A are **interior** and **closure** operations in the topological space T

2.4 Approximations and membership relation

$$x \underline{\in}_R X \text{if and only if } x \in \underline{R}X$$
$$x \overline{\in}_R X \text{if and only if } x \in \overline{R}X$$

where $\underline{\in}_R$ read "x **surely belongs** to X w.r.t. R" and $\overline{\in}_R$ - "x **possibly belongs** to X w.r.t. R". The **lower** and **upper** membership.

Proposition 2.3. 1. *$x \underline{\in} X$ implies $x \in X$ implies $x \overline{\in} X$*

2. *$X \subset Y$ implies ($x \underline{\in} X$ implies $x \underline{\in} Y$ and $x \overline{\in} X$ implies $x \overline{\in} Y$)*

3. *$x \overline{\in} (X \cup Y)$ if and only if $x \overline{\in} X$ or $x \overline{\in} Y$*

4. *$x \underline{\in} (X \cap Y)$ if and only if $x \underline{\in} X$ and $x \underline{\in} Y$*

5. *$x \underline{\in} X$ or $x \underline{\in} Y$ implies $x \underline{\in} (X \cup Y)$*

6. *$x \overline{\in} X \cap Y$ implies $x \overline{\in} X$ and $x \overline{\in} Y$*

7. *$x \underline{\in} (-X)$ if and only if non $x \overline{\in} X$*

8. *$x \overline{\in} (-X)$ if and only if non $x \underline{\in} X$*

2.5 Numerical characterization of imprecision

accuracy measure

$$\alpha_R(X) = \frac{card\ \underline{R}}{card\ \overline{R}}$$

2.6 Topological characterization of imprecision

Definition 2.1. 1. *If $\underline{R}X \neq \emptyset$ and $\overline{R}X \neq U$, then we say that X is **roughly R-definable**. We can decide whether some elements belong to X or $-X$*

2. *If $\underline{R}X = \emptyset$ and $\overline{R}X \neq U$, then we say that X is **internally R-undefinable**. We can decide whether some elemnts belong to $-X$*

3. *If $\underline{R}X \neq \emptyset$ and $\overline{R}X = U$, then we say that X is **externally R-undefinable**. We can decide whether some elements belong to X*

4. *If $\underline{R}X = \emptyset$ and $\overline{R}X = U$, then we say that X is **totally R-undefinable**. unable to decide*

Proposition 2.4 (2.4)**.** 1. *Set X is R-definable(roughly R-definable, totally R-undefinable) if and only if so is $-X$*

2. *Set X is externally R-undefinable if and only if $-X$ is internally R-undefinable*

Proof. 1.

$$\begin{aligned} R\text{-definable} &\Leftrightarrow \underline{R}X = \overline{R}X, \underline{R} \neq \emptyset, \overline{R} \neq U \\ &\Leftrightarrow -\underline{R}X = -\overline{R}X \\ &\Leftrightarrow \overline{R}(-X) = \underline{R}(-X) \end{aligned}$$

□