IEEE SMC · NAFIPS North American Fuzzy Information Processing Society · IEEE

NAFIPS 2004

2004 Annual Meeting of the North American Fuzzy Information Processing Society
Banff, Alberta, Canada

FUZZY SETS IN THE HEART OF THE CANADIAN ROCKIES

JUNE 27 - 30, 2004

ISBN 0-7803-8377-X IEEE Cat. No. 04TH8736C

Edited by Scott Dick, Lukasz Kurgan, Petr Musilek, Witold Pedrycz, and Marek Reformat

technical support: nafips04@ece.ualberta.ca

Approximating Functions Using Rough Sets

Zdzisław Pawlak
Institute for Theoretical and Applied Informatics
Polish Academy of Sciences
Bałtycka 5, 44-100 Gliwice, Poland
and
Warsaw School of Information Technology*
Newelska 6, 01-447 Warsaw, Poland
zpw@ii.pw.edu.pl

James Peters
Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba R3T 5V6, Canada
jfpeters@ee.umanitoba.ca

Andrzej Skowron
Institute of Mathematics, Warsaw University
Banacha 2, 02-097 Warsaw, Poland
skowron@mimuw.edu.pl

Abstract

Approximating of functions that are specified using imperfect knowledge is one of the central issues of many areas such as machine learning, pattern recognition, data mining, or qualitative reasoning. However, we do not have yet satisfactory methods for approximation of functions and developed calculi on function approximations. In the paper we discuss a function approximation using the rough set approach. The main difference with the existing approaches in rough set theory is based on modification of the inclusion measure. This makes it possible to overcome some drawbacks of the previously used definitions. For applications it is important to develop rough measures on approximated objects, in particular on function approximations. The modified inclusion measure is also used to define an exemplary measure, i.e., the rough integral.

Keywords: Approximation, function, rough sets, rough inclusion, rough measures, rough integral

*former University of Information Technology and Management

1 Introduction

Approximating of functions specified by imperfect knowledge is one of the central issues in many areas such as machine learning, pattern recognition, data mining, or qualitative reasoning (see, e.g., [1, 2, 3, 23]). Moreover, methods for function approximation are important for real-life applications, e.g., in control or image processing (see, e.g., [7, 24, 6]).

The approximation of objects is one of the central issues in rough set theory and rough mereology (see, e.g., [9, 28, 21]) as well as in granular computing. There have been already reported results on function approximation in the literature on rough sets (see, e.g., [8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20]). A function is approximated by means of knowledge granules defined by indiscernibility neighborhoods and a rough inclusion function making it possible to measure degrees to which neighborhoods are matching the function. However, using the existing approaches based on the standard inclusion measure of sets in Cartesian products the lower approximation of functions is, in almost all cases, equal to the empty set. We define an inclusion measure more relevant for a function approximation. In particular, this inclusion measure does not lead to the above mentioned drawback.

To measure the quality of approximation and to reason about approximated objects some new special measures are constructed. We introduce an example of such a measure defined by rough integrals of samples of real functions. In the definition the introduced inclusion measure is used.

2 Approximation spaces

Let us recall the definition of an approximation space from [26, 9]. For simplicity of reasoning we omit parameters that label components of approximation spaces.

An *approximation space* is a system $AS = (U, I, \nu)$, where

- U is a non-empty finite set of objects,
- $I : U \to P(U)$ is a partial function, called the uncertainty function, such that $x \in I(x)$ for any x from the domain of I,
- $\nu : P(U) \times P(U) \to [0, 1]$ is a rough inclusion function.

A set $X \subseteq U$ is *definable in AS* if and only if it is a union of some values of the uncertainty function.

The standard rough inclusion function ν_{SRI} defines the degree of inclusion between two subsets of U by

$$\nu_{SRI}(X, Y) = \begin{cases} \frac{card(X \cap Y)}{card(X)} & \text{if } X \neq \emptyset \\ 1 & \text{if } X = \emptyset. \end{cases} \quad (1)$$

The lower and the upper approximations of subsets of U are defined as follows.

For any approximation space $AS = (U, I, \nu)$, $0 \leq p < q \leq 1$, and any subset $X \subseteq U$ the q-lower and the p-upper approximation of X in AS are defined by

$$LOW_q(AS, X) = \{x \in U : \nu(I(x), X) \geq q\}, \quad (2)$$

$$UPP_p(AS, X) = \{x \in U : \nu(I(x), X) > p\}, \quad (3)$$

respectively.

Then the boundary region of X in AS is defined by

$$BN_{p,q}(AS, X) = UPP_p(AS, X) - LOW_q(AS, X). \quad (4)$$

Approximation spaces can be constructed directly from information systems or from information systems enriched by some similarity relations on attribute value vectors. The above definition generalizes several approaches existing in literature like those based on an equivalence or tolerance indiscernibility relation as well as those based on exact inclusion of indiscernibility classes into concepts [9, 29]. Such inclusion measures have been widely used for years by data mining and rough set communities. However, Jan Łukasiewicz [5] was the first who used this idea to estimate the probability of implications.

One can observe that it is possible to generalize the discussed approximation spaces by considering a family of neighborhoods covering the object space U instead of the uncertainty function I [28].

3 Relation and function approximation

One can directly apply the definition of set approximation to relations. For simplicity, but without loss of generality, we consider binary relations only. Let $R \subseteq U \times U$ be a binary relation. We can consider approximation of R by an approximation space $AS = (U \times U, I, \nu)$ in an analogous way as in Section 2:

$$LOW_q(AS, R) = \{(x, y) \in U \times U : \nu(I(x, y), R) \geq q\} \quad (5)$$

$$UPP_p(AS, R) = \{(x, y) \in U \times U : \nu(I(x, y), R) > p\} \quad (6)$$

for $0 \leq p < q \leq 1$. The main problem is how to construct relevant approximation spaces, i.e., how to define uncertainty and inclusion functions. One of the solutions is the following uncertainty function

$$I(x, y) = I(x) \times I(y), \quad (7)$$

(assuming that the one dimensional uncertainty function is given) and the standard inclusion, i.e., $\nu = \nu_{SRI}$.

Now, let us consider an approximation space $AS = (U, I, \nu)$ and a function $f : Dom \longrightarrow U$, where $Dom \subseteq U$. By $Graph(f)$ we denote the set $\{(x, f(x)) : x \in Dom\}$. We can easily see that if we apply the above definition of relation approximation to f (it is a special case of relation) then the lower approximation is almost always empty. Thus, we have to construct the relevant approximation space $AS^* = (U \times U, I^*, \nu^*)$ in different way, e.g., by extending the uncertainty function as well as the inclusion function on subsets of $U \times U$. We assume that the value $I^*(x, y)$ of the uncertainly function, called the neighborhood (or the window) of (x, y) is defined by

$$I^*(x, y) = I(x) \times I(y), \quad (8)$$

for $(x, y) \in U \times U$.

Next, we should decide how to define values of the inclusion function on pairs $(I^*(x, y), Graph(f))$, i.e., how to define the degree r to which the intersection $I^*(x, y) \cap Graph(f)$ is included into the $Graph(f)$.

One can consider a ratio

$$r = \frac{card(\{x \in I(x) \cap Dom : f(x) \in I(y)\})}{card(I(x))}, \quad (9)$$

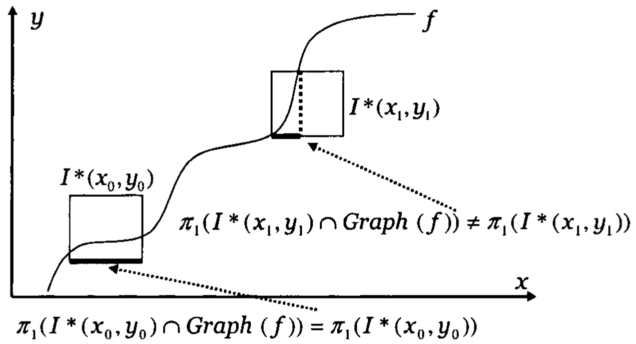

Figure 1. Function approximation

i.e., the ratio of the number of all objects from $I(x) \cap Dom$ on which f takes a value from $I(y)$ to the number of all objects in $I(x)$.

If $r = 1$ then (x, y) defining the window $I^*(x, y)$ is in the lower approximation of $Graph(f)$.

If $0 < r \leq 1$ then (x, y) defining the window $I^*(x, y)$ is in the upper approximation of $Graph(f)$.

Using the above intuition, we assume that the inclusion holds to degree one if the domain of $Graph(f)$ restricted to $I(x)$ is equal to $I(x)$. This can be formally defined by the following condition:

$$\pi_1(I^*(x,y) \cap Graph(f)) = \pi_1(I^*(x,y)) \qquad (10)$$

where π_1 denotes the projection on the first coordinate. Condition (10) is equivalent to:

$$I(x) \subseteq Dom \text{ and } f(I(x)) \subseteq I(y). \qquad (11)$$

Thus, the inclusion function (called *rough inclusion*) ν^* for subsets $X, Y \subseteq U \times U$ is defined by

$$\nu^*(X,Y) = \begin{cases} \frac{card(\pi_1(X \cap Y))}{card(\pi_1(X))} & \text{if } \pi_1(X) \neq \emptyset \\ 1 & \text{if } \pi_1(X) = \emptyset. \end{cases} \qquad (12)$$

Hence, the relevant inclusion function in approximation spaces for function approximations is such a function that does not measure the degree of inclusion of its arguments but their perceptions, represented in the above example by projections of corresponding sets. Certainly, one can chose another definition based, e.g., on the density of pixels (in case of images) in the window that are matched by the function graph.

We have the following proposition:

Proposition 1. Let $AS^* = (U \times U, I^*, \nu^*)$ be an approximation space with I^*, ν^* defined by (8), (12), respectively, and let $f : Dom \longrightarrow U$ where $Dom \subseteq U$. Then we have

1. $(x, y) \in LOW_1(AS^*, Graph(f))$ if and only if $f(I(x)) \subseteq I(y)$ and $I(x) \subseteq Dom$,
2. $(x, y) \in UPP_0(AS^*, Graph(f))$ if and only if $f(I(x)) \cap I(y) \neq \emptyset$.

In case of arbitrary parameters p, q satisfying $0 \leq p < q \leq 1$ we have

Proposition 2. Let $AS^* = (U \times U, I^*, \nu^*)$ be an approximation space with I^*, ν^* defined by (8), (12), respectively, and let $f : Dom \longrightarrow U$ where $Dom \subseteq U$. Then we have

1. $(x, y) \in LOW_q(AS^*, Graph(f))$ if and only if $card(\{x' \in I(x) \cap Dom : f(x') \in I(y)\}) \geq q \cdot card(I(x))$,
2. $(x, y) \in UPP_p(AS^*, Graph(f))$ if and only if $card(\{x' \in I(x) \cap Dom : f(x') \in I(y)\}) > p \cdot card(I(x))$.

From the propositions it follows, e.g., that for any $x_1 \in Dom$ if the pair $(x_1, f(x_1))$ belongs to the lower approximation of $Graph(f)$ then the graph of f restricted to $I(x_1)$ in completely included into the window $I^*(x_1, f(x_1))$.

In our example, we have defined the inclusion degree between two subsets of Cartesian product using the inclusion degree between their projections.

4 An Example: Rough Integral

In this section we present an example showing how to use the rough inclusion (12) to define a rough measure, called rough integral.

For simplicity of reasoning we consider a function $f : Dom \longrightarrow U$ where $Dom \subseteq U$ and U is a finite subset (called a sample) of the set R_+ of non-negative reals. Such functions can be treated as samples of functions from R_+ into R_+. We also restrict our discussion to the approximation of such samples, i.e., the approximation problems of extensions of such samples will be discussed elswhere using the rough set approach to concept approximation presented in [28].

Let $\mathcal{O}$ be a given family of neighborhoods $I \subseteq R^2$ that are intervals in R^2. For any $I \in \mathcal{O}$ by I_1, I_2 we denote projections $\pi_1(I), \pi_2(I)$ of I on the axes.

For a given (finite) sample $U \subset R_+$ we will also consider neighborhoods obtained from $\mathcal{O}$ by their restrictions to U, i.e., neighborhoods of the form $I \cap (U \times U)$ (see Figure 2) for $I \in \mathcal{O}$. Such neighborhoods will be used to approximate real valued functions restricted to a given (finite) sample U (called samples of real valued functions, for short). Assuming that $f : Dom \longrightarrow U$ is a sample of a real valued function one can distinguish a family consisting of subfamiles of $\mathcal{O}$ "matching" $Graph(f)$, i.e.,

$$Fam_f = \{\mathcal{F} \subseteq \mathcal{O} : (\mathcal{F} \text{ matches } Graph(f)) \wedge \forall I, I' \in \mathcal{F}(I \neq I' \Rightarrow I \cap I' = \emptyset)\}. \quad (13)$$

From the definition it follows that for any $\mathcal{F} \in Fam_f$ and $(x, y) \in U \times U$ there exists at most one $I \in \mathcal{F}$ such that $(x, y) \in I$. Hence, one can define a (partial) uncertainty function I_U assigning for any $(x, y) \in U \times U$ the neighborhood

$$I_U(x, y) = I \cap (U \times U), \quad (14)$$

where $(x, y) \in I$, if such neighborhood $I \in \mathcal{F}$ exists.

One can also observe that the phrase "$\mathcal{F}$ matches $Graph(f)$" can be easily defined by the boundary region (i.e., the difference between the upper approximation and the lower approximation) of $Graph(f)$ in the approximation space defined by neighborhoods defined from $\mathcal{F}$ by I_U (see (14)) and the rough inclusion (12). The matching can be exact (to degree one), if the boundary region is empty or $Graph(f)$ can be matched to a degree that can be expressed by the (relative) boundary region size. Certainly, we use the rough inclusion (12) in definitions of the lower approximation (2), the upper approximation (3), and the boundary region (4). For example, "$\mathcal{F}$ matches exactly $Graph(f)$" if and only if

$$LOW(AS_{\mathcal{F}}, Graph(f)) = Graph(f), \quad (15)$$

where

$$AS_{\mathcal{F}} = (U \times U, I_U, \nu). \quad (16)$$

In (16) I_U is the uncertainty function defined by (14), and ν is the rough inclusion (12).

From the above discussion it follows that an additional parameter of Fam_f should be chosen that expresses the degree to which the lower approximation of f defined by any family from Fam_f should approximate $Graph(f)$.

For any $\mathcal{F} \in Fam_f$ we define the $\mathcal{F}$-lower sum and $\mathcal{F}$-upper sums of f by

$$Low(\mathcal{F}, f) = \sum_{I \in \mathcal{F}} length(I_1) \cdot min(I_2), \quad (17)$$
$$Upp(\mathcal{F}, f) = \sum_{I \in \mathcal{F}} length(I_1) \cdot max(I_2),$$

respectively, where $length(I_1) = max\{y - x : x, y \in I_1\}$.

Assume that

$$Quality(\mathcal{F}, f) = Upp(\mathcal{F}, f) - Low(\mathcal{F}, f). \quad (18)$$

We can now define the rough integral (optimization) problem.

ROUGH INTEGRAL PROBLEM (RI)
INPUT: $f : Dom \longrightarrow U$, $\mathcal{O}$ and $\varepsilon \in (0, 1)$;
OUTPUT:

- (a, b) if there exists $\mathcal{F}_{opt} \in Fam_f$ such that the following conditions are satisfied:

1. $|\mathcal{F}_{opt}| < \varepsilon|\mathcal{O}|$,
2. $Quality(\mathcal{F}_{opt}, f) = min_{\mathcal{F} \in Fam_f} Quality(\mathcal{F}, f)$,
3. $(a, b) = (Upp(\mathcal{F}_{opt}, f), Low(\mathcal{F}_{opt}, f))$;

- NONE, otherwise.

The first condition in the RI problem is expressing the requirement that the considered families of neighborhoods used to match the graph of the function should be "small" relative to $\mathcal{O}$. In the second condition a subfamily of neighborhoods minimizing the difference between the upper sum and the lower sum (among all considered subfamilies) is chosen.

In general, the solution of the RI problem is not unique (one can obtain different intervals with the same length). However, the set of solutions (if it is nonempty) can be treated in some cases as unique to a degree. To explain this let us consider first a simple definition.

We call two intervals (a, b), (c, d) of reals *δ-close*, where $\delta > 0$ is a given threshold if and only if

$$\frac{|(a, b) \div (c, d)|}{|(a, b) \cup (c, d)|} < \delta \quad (19)$$

where $\div$ is the set theoretical symmetric difference of sets.

Now, assuming that δ is a given threshold we say that the *rough integral* of f relative to $\mathcal{O}$ and ε is defined with precision δ if and only if the set of solutions of the RI problem (with input f, $\mathcal{O}$ and ε) is nonempty and any two its solutions are δ-close. If the rough integral of f (relative to $\mathcal{O}$ and ε) is defined to a degree δ then any solution of the RI problem is called the *rough integral* of f (relatively to $\mathcal{O}$ and ε) with precision δ.

Observe that the rough integral has been defined here using a bounded discernibility between reals contrary to the classical definition of integral of real functions (see, e.g., [25]).

Investigating such measures is important, e.g., for approximate or qualitative reasoning in control systems based on the rough set approach. For example, some rough integrals were used for localization of objects and for selection of relevant sensors (see, e.g., [20, 22]).

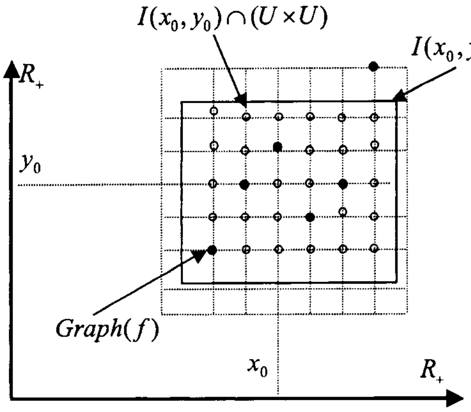

Figure 2. Neighborhood $I(x_0, y_0) \cap (U \times U)$

5 Conclusions

We have discussed approximation of functions specified by imperfect knowledge. We also presented rough integrals for samples of real functions as an example of rough measures. We plan to extend the approach to approximation of real functions specified by their samples [23]. Our approach will be based on the rough set approach to concept approximation [28] and combination of rough set theory with the case-based reasoning approach [4].

The main goal of our project is also to develop calculi for approximations of functions defined on information granules with values that are information granules too [18, 27].

Acknowledgements

The research of James Peters has been supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) 185986.

The research of Andrzej Skowron has been supported by the grant 3 T11C 002 26 from Ministry of Scientific Research and Information Technology of the Republic of Poland.

References

[1] J. Friedman, T. Hastie, and R. Tibshirani. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag, Heidelberg, Germany, 2001.

[2] W. Kloesgen and J. Żytkow, editors. *Handbook of Knowledge Discovery and Data Mining.* Oxford University Press, Oxford, 2002.

[3] B. Kuipers. Qualitative simulation. In R. A. Mayers, editor, *Encyclopedia of Physical Science and Technology, Third Edition*, pages 287–300. Academic Press, New York, 2001.

[4] D. B. Leake. *Case-Based Reasoning: Expriences, Lessons, & Future Directions.* AAAI Press, Menlo Park, CA, 1996.

[5] J. Łukasiewicz. Die logischen grundlagen der wahrscheinilchkeitsrechnung, Kraków 1913. In L. Borkowski, editor, *Jan Łukasiewicz - Selected Works.* North Holland Publishing Company, Amstardam, London, Polish Scientific Publishers, Warsaw, 1970.

[6] A. Nakamura and A. Rosenfeld. Digital calculus. *Information Sciences*, 98:63–98, 1997.

[7] M. Papa, H. M. Tai, and S. Shenoi. Design and evaluation of control systems using cell mapping techniques. *IEEE Control Systems*, 17:52–65, 1997.

[8] Z. Pawlak. Rough functions. *Bull. Polish Acad. Sci. Tech.*, 35(5-6):249–251, 1987.

[9] Z. Pawlak. *Rough Sets: Theoretical Aspects of Reasoning about Data*, volume 9 of *System Theory, Knowledge Engineering and Problem Solving*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[10] Z. Pawlak. *Rough Sets, Rough Real Functions*, volume 50 of *ICS WUT Reports.* Institute of Computer Science, Warsaw University of Technology, Warsaw, 1994.

[11] Z. Pawlak. *On Rough Derivatives, Rough Integrals and Rough Differential Equations*, volume 41 of *ICS WUT Reports.* Institute of Computer Science, Warsaw University of Technology, Warsaw, 1995.

[12] Z. Pawlak. *On Some Issues Connected with Roughly Continuous Functions*, volume 21 of *ICS WUT Reports.* Institute of Computer Science, Warsaw University of Technology, Warsaw, 1995.

[13] Z. Pawlak. *Rough Calculus*, volume 58 of *ICS WUT Reports.* Institute of Computer Science, Warsaw University of Technology, Warsaw, 1995. (see also abstract in Proceedings of the 2nd Annual Joint Conference on Information Science (1995), 34-35).

[14] Z. Pawlak. *Rough Real Functions and Rough Controllers*, volume 1 of *ICS WUT Reports.* Institute of Computer Science, Warsaw University of Technology, Warsaw, 1995.

[15] Z. Pawlak. Rough sets, rough relations and rough functions. *Fundamenta Informaticae*, 27(2-3):103–108, 1996.

[16] Z. Pawlak. Rough real functions and rough controllers. In T. Y. Lin and N. Cercone, editors, *Rough Sets and Data Mining: Analysis of Imprecise Data*, pages 139–147. Kluwer Academic Publishers, Boston, Dordrecht, 1997.
[17] Z. Pawlak. Rough sets, rough functions and rough calculus. In S. K. Pal and A. Skowron, editors, *Rough Fuzzy Hybridization: A New Trend in Decision Making*, pages 99–109. Springer Verlag, Singapore, 1999.
[18] Z. Pawlak, J. Peters, and A. Skowron. A rough set approach to measuring information granules. In *IEEE COMPSAC'02 Workshop on Data Mining and Rough Computing, August 2002*, pages 1135–1139, Oxford, GB, 2002.
[19] Z. Pawlak, J. Peters, A. Skowron, Z. Suraj, S. Ramanna, and M. Borkowski. Rough measures and integrals: A brief introduction. In T. Terano, T. Nishida, A. Namatame, S. Tsumoto, Y. Ohsawa, and T. Washio, editors, *New Frontiers in Artificial Intelligence, Joint JSAI'01 Workshop Post-Proceedings*, volume 2253 of *Lecture Notes in Artificial Intelligence*, pages 374–379, Heidelberg, 2001. Springer-Verlag.
[20] Z. Pawlak, J. Peters, A. Skowron, Z. Suraj, S. Ramanna, and M. Borkowski. Rough measures, rough integrals and sensor fusion. In M. Inuiguchi, S. Hirano, and S. Tsumoto, editors, *Rough Set Theory and Granular Computing*, volume 125 of *Studies in Fuzziness and Soft Computing*, pages 263–272. Springer-Verlag, Heidelberg, 2003.
[21] Z. Pawlak, L. Polkowski, and A. Skowron. Rough set theory. In B. Wah, editor, *Encyclopedia of Computer Science and Engineering*. Wiley, New York, USA, 2004. (to appear).
[22] J. F. Peters, S. Ramanna, M. Borkowski, A. Skowron, and Z. Suraj. Sensor, filter, and fusion models with rough petri nets. *Fundamenta Informatica*, 47(3/4):307–323, 2001.
[23] T. Poggio and S. Smale. The mathematics of learning: Dealing with data. *Notices of the AMS*, 50(5):537–544, 2003.
[24] A. Rosenfeld. Continuouis functions on digital pictures. *Pattern Recognition Letters*, 4:177–184, 1986.
[25] S. Saks. *Theory of the Integral. With Notes by S. Banach.* Dover Publications, New York, 1994.
[26] A. Skowron and J. Stepaniuk. Tolerance approximation spaces. *Fundamenta Informaticae*, 27(2-3):245–253, 1996.
[27] A. Skowron and J. Stepaniuk. Information granules and rough-neural computing. In S. K. Pal, L. Polkowski, and A. Skowron, editors, *Rough-Neural Computing: Techniques for Computing with Words*, Cognitive Technologies, pages 43–84. Springer-Verlag, Heidelberg, Germany, 2003.
[28] A. Skowron, R. Swiniarski, and P. Synak. Approximation spaces and information granulation (to appear). In *Fourth International Conference on Rough Sets and Current Trends in Computing (RSCTC'04), Uppsala, Sweden, June 1-5, 2004*, Lecture Notes in Computer Science. Springer-Verlag, Heidelberg, Germany, 2004.
[29] W. Ziarko. Variable precision rough set model. *Journal of Computer and System Sciences*, 46:39–59, 1993.