Ryszard Janicki is a professor at McMaster University, Hamilton, Ontario, Canada in the Department of Computing and Software. He received the M.Sc. degree in Applied Mathematics from the Warsaw University of Technology, Poland in 1975, and the Ph.D. and Habilitation in Computer Science from the Polish Academy of Sciences, Warsaw, Poland in 1977 and 1981 respectively.

He worked at Warsaw University of Technology, Warsaw, Poland (1975-1984), Aalborg University, Aalborg, Denmark (1984-1986) and, since 1986, at McMaster University.

Dr. Janicki published about two hundred papers and two monographs. His research interests include concurrency theory, abstract approximation, rough sets, ranking theory, fundamentals of software engineering and mereology.

He promoted sixteen PhD students.

On Optimal Approximations for Rough Sets

Ryszard Janicki

Department of Computing and Software,
McMaster University,
Hamilton, ON, L8S 4K1 Canada
janicki@mcmaster.ca

Classical Rough Sets [7,8] 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. 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*) [1,9], and there are some techniques for finding *maximal similarity* (*minimal distance*) between entities and their approximations.

In [4,5] the concept of *optimal approximation* has been added to standard Rough Sets and discussed in some detail, and in [3] it was extended to the Rough Sets defined by coverings.

The word ‘optimal’ suggests some numerical calculations and comparisons; we need some proper definition of *similarity measure* for sets. We will discuss several such definitions, propose and justify some set of axioms for the similarity concept, as well the concept of *consistency*, that allows treating some similarities as equivalent.

An efficient greedy algorithm for finding optimal approximations for standard Rough Sets (i.e. induced by partitions) and *Marczewski-Steinhaus similarity measure* [6] (and all similarity measures consistent with it) will be discussed in detail. The algorithm is based on the properties of an index that quantifies the ratio of common to distinct elements of two given sets. We used the Marczewski-Steinhaus similarity measure as an engine of our algorithm because it has a very natural and regular definition and convenient mathematical properties. Moreover, the Marczewski-Steinhaus similarity measure is a generalization of a very popular Jaccard index [2], and it is also consistent (i.e. practically equivalent) to many other popular similarity measures.

In [3] a fairly natural transition from Rough Sets induced by coverings to standard Rough Sets (i.e. induced by partitions) have been proposed and discussed. We will use this transformation to extend our algorithm for Rough Sets induced by coverings.

References

1. Deza M.M., Deza E., *Encyclopedia of Distances*, Springer, Berlin, 2012.
2. Jaccard P., Étude comparative de la distribution florale dans une portion des Alpes et des Jura, *Bulletin de la Société Vaudoise des Sciences Naturalles* 37 (1901) 547-549.
3. Janicki R., Yet Another Kind of Rough Sets Induced by Coverings, *Lecture Notes in Artificial Intelligence* 10313, Springer 2017, 140-153

4. Janicki R., Lenarčič A., Optimal Approximations with Rough Sets, *Lecture Notes in Artificial Intelligence* 8171, Springer 2013, 87–98.
5. Janicki R., Lenarčič A., Optimal Approximations with Rough Sets and Similarities in Measure Spaces, *International Journal of Approximate Reasoning*, 71 (2016) 1–14.
6. Marczewski E., Steinhaus H., On a certain distance of sets and corresponding distance of functions, *Colloquium Mathematicum*, 4 (1958), 319–327.
7. Pawlak Z., Rought Sets, *International Journal of Computer and Information Sciences*, 34 (1982) 557–590.
8. Pawlak Z., *Rough Sets*, Kluwer, Dordrecht 1991.
9. Tversky A., Features of similarity, *Psychological Reviews* 84, 4 (1977) 327–352.