ADVANCES IN *SOFT COMPUTING*

Advances in Soft Computing

Lech Polkowski

Rough Sets

Mathematical Foundations

Springer-Verlag Berlin Heidelberg GmbH

ADVANCES IN SOFT COMPUTING

Advances in Soft Computing

Lech Polkowski

Rough Sets

Mathematical Foundations

Springer-Verlag Berlin Heidelberg GmbH

Rough Sets

Advances in Soft Computing

Editor-in-chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw, Poland
E-mail: kacprzyk@ibspan.waw.pl
http://www.springer.de/cgi-bin/search-bock.pl?series=4240

Esko Turunen
Mathematics Behind Fuzzy Logic
1999. ISBN 3-7908-1221-8

Robert Fullér
Introduction to Neuro-Fuzzy Systems
2000. ISBN 3-7908-1256-0

Robert John and Ralph Birkenhead (Eds.)
Soft Computing Techniques and Applications
2000. ISBN 3-7908-1257-9

Mieczysław Kłopotek, Maciej Michalewicz
and Sławomir T. Wierzchoń (Eds.)
Intelligent Information Systems
2000. ISBN 3-7908-1309-5

Peter Sinčák, Ján Vaščák, Vladimír Kvasnička
and Radko Mesiar (Eds.)
The State of the Art in Computational Intelligence
2000. ISBN 3-7908-1322-2

Bernd Reusch and Karl-Heinz Temme (Eds.)
Computational Intelligence in Theory and Practice
2001. ISBN 3-7908-1357-5

Robert John and Ralph Birkenhead (Eds.)
Developments in Soft Computing
2001. ISBN 3-7908-1361-3

Mieczysław A. Kłopotek, Maciej Michalewicz
and Sławomir T. Wierzchoń (Eds.)
Intelligent Information Systems 2001
2001. ISBN 3-7908-1407-5

Antonio Di Nola and Giangiacomo Gerla (Eds.)
Lectures on Soft Computing and Fuzzy Logic
2001. ISBN 3-7908-1396-6

Tadeusz Trzaskalik and Jerzy Michnik (Eds.)
Multiple Objective and Goal Programming
2002. ISBN 3-7908-1409-1

James J. Buckley and Esfandiar Eslami
An Introduction to Fuzzy Logic and Fuzzy Sets
2002. ISBN 3-7908-1447-4

Ajith Abraham and Mario Köppen (Eds.)
Hybrid Information Systems
2002. ISBN 3-7908-1480-6

Mieczysław A. Kłopotek, Sławomir T. Wierzchoń
and Maciej Michalewicz (Eds.)
Intelligent Information Systems 2002
2002. ISBN 3-7908-1509-8

Lech Polkowski

Rough Sets

Mathematical Foundations

With 20 Figures

Springer-Verlag Berlin Heidelberg GmbH

Prof. Dr. Lech Polkowski

Polish-Japanese Institute of Information Technology
ul. Koszykowa 86
02-008 Warsaw
Poland

and

Department of Mathematics and Computer Science
University of Wormia and Mazury
ul. Żofnierska 14a
10-561 Olsztyn
Poland
Lech.Polkowski@pjwstk.edu.pl

ISSN 1615-3871
ISBN 978-3-7908-1510-8 ISBN 978-3-7908-1776-8 (eBook)
DOI 10.1007/978-3-7908-1776-8

Library of Congress Cataloging-in-Publication Data applied for
Die Deutsche Bibliothek – CIP-Einheitsaufnahme
Polkowski, Lech: Rough sets: mathematical foundations / Lech Polkowski. – Heidelberg; New York: Physica-Verl., 2002
(Advances in soft computing)
ISBN 978-3-7908-1510-8

Originally published by Physica-Verlag Heidelberg New York in 2002

Softcover Design: Erich Kirchner, Heidelberg

SPIN 10882886 88/2202-5 4 3 2 1 0 – Printed on acid-free paper

Dedicated to Professor Zdżisław Pawlak on his 75th birthday

Preface

Er wunderte sich, dass den Katzen gerade an der Stelle zwei Löcher in den Pelz geschnitten werden, wo sie die Augen hätten

Georg Christoph Lichtenberg, *Aphorismen*

Rough sets, similarly to fuzzy sets, have been invented to cope with uncertainty. Both paradigms address the phenomenon of non – crisp concepts, notions which, according to Frege, are characterized by a presence of a non – empty *boundary* which does encompass objects neither belonging with certainty to the given concept nor belonging with certainty to its complement.

A crisp concept $A \subseteq X$ induces on the universe X the partition $\{A, X \setminus A\}$ into objects belonging to A and objects belonging to $X \setminus A$ as the problem whether $x \in A$ is decidable on the basis of knowledge available to us.

The partition $\{A, X \setminus A\}$ can be rendered in functional language into the *characteristic function* χ_A which assumes the value 1 on A and the value 0 on $X \setminus A$.

Assume our knowledge increases due to our experience, new experiments etc. etc. In the language of partitions, this results in passing from the partition $\{A, X \setminus A\}$ to a partition $\mathcal{P}$ whose members are subsets of the universe containing objects indiscernible one to another on the basis of new knowledge.

In consequence, the formerly crisp notion A may become non – crisp: for some objects x the problem whether $x \in A$ may become undecidable: some objects indiscernible to x may be in A while some objects indiscernible to x may be not in A.

Either paradigm solves the description problem of A in its own way; rough sets approach this problem by considering *approximations* to A induced by the partition $\mathcal{P}$ viz. into the *lower approximation* A_P those objects fall whose whole class $[x]_P$ is contained in A. We can by all means conclude that such objects belong in A with *certainty* (or, *necessarily*). To the *upper approximation* A^P those objects fall whose class $[x]_P$ intersects A but is not contained in A; we can conclude that such objects belong *possibly* in A. The *boundary*

$B_P = A^P \setminus A_P$ consists of ambiguous objects possibly but not certainly in A.

Fuzzy sets approach the problem of description of A in the other way: the characteristic function $\chi_A : X \to \{0,1\}$ undergoes a modification to a *fuzzy characteristic* function $\mu_A : X \to [0,1]$. We can regard objects in the inverse $\mu_A^{-1}(1)$ as belonging in A with certainty while objects in $\mu_A^{-1}((0,1])$ belong in A possibly.

Both approaches are related; assuming e.g. the case of a finite universe X, and a partition $\mathcal{P}$, we can assign to each $x \in X$, a quotient $\mu_P(x) = \frac{|A\cap[x]_P|}{|[x]_P|}$ according to an idea of Pawlak and Skowron. The function μ_P is then a fuzzy characteristic function inducing same approximations as the partition P. Conversely, given a fuzzy characteristic function χ_A, we can induce a partition P_χ by letting

$$[x]_{P_\chi} = [y]_{P_\chi} \Leftrightarrow \chi_A(x) = \chi_A(y).$$

Then certainty and possibility regions induced by χ_A correspond to the lower and upper approximations induced by P_χ.

From this point both paradigms go their own ways. It is the aim of this book to present mathematical foundations and methods of the rough set theory. As one may expect, rough set theory with its notions of the lower approximation, the upper approximation and the boundary region is related to 3 – valued logics, while in fuzzy set theory various infinite valued logics are reflected. The topic of mutual relationships and influences of each to other is taken up in the end (cf. Chapter 14 where we outline the ideas of *rough fuzzy* and *fuzzy rough* sets leading to hybrid constructions).

Mathematical foundations of rough set theory are constructed from fragments of set theory, logic, topology and algebra. We aim at presenting a full picture of results obtained in rough set theory by set–theoretical, logical, topological and algebraic methods.

We intend at making a complete and self – contained exposition of the necessary material, so in addition to advanced topics presented in Chapters 10 – 12 on rough set – theoretic results, we include all necessary preliminaries from the concerned areas of mathematics so the reader may find in the book along with results on rough sets also all information from the relevant mathematical field. In accordance with this desire, we include as an interlude between Chapter 1 bringing the basic information about rough sets and intended as a self–contained tutorial, and Chapter 10, a series of Chapters devoted to respective areas of pure mathematics. Chapter 1 contains moreover a bibliography of 420 recent papers in rough set theory.

The reader will find in Chapter 2 a short but precise and rigorous course in the sentential logic covering the classical deductive theory as well as the Gentzen – style axiomatization in the version due to Kanger. Chapter 3 brings forth an exposition of Syllogistic: in the light of the Słupecki result on Syllogistic as the complete logical theory of containment and intersection i.e. the basic tools in defining rough set approximations, Syllogistic appears as one of the main logical reflexions of rough set theory.

In Chapter 4, many – valued finite – valued logics are encountered for the first time. We discuss the 3 – valued logic of Łukasiewicz and we introduce the Wajsberg axiomatization of this logic along with the proof of its completeness. Many–valued logics are discussed in the Rosser – Turquette spirit and finally 4 – valued logic of Łukasiewicz is introduced as a vehicle to present the basic modal logic.

Modal logics are discussed in Chapter 5 in which we also introduce the Kripke semantics and prove the completeness of basic logics: K, T, S4, and S5.

Set theory is exposed in Chapter 6. In addition to standard notions of naive set theory recalled, we discuss basic constructions of formal set theory among them Cartesian products, relations, orderings, and filters and essentially important notions of a lattice and of a Boolean algebra. A special attention is devoted to equivalence and tolerance relations as to constructs rough set theory deals with frequently and substantially.

Chapter 7 is devoted to topological structures and we discuss basic notions and most important results of set theoretic topology among them compactness and completeness. Important topics like regular sets, filters, Stone representation, are brought in there. As a preliminary to a discussion of rough set topologies in Chapter 11, we include a section on topologies on collections of closed sets containing a discussion of the Hausdorff – Pompéju metric.

A discussion of basic notions of lattice theory is carried out in Chapter 8. We highlight the notions of a distributive lattice, a Stone algebra, pseudo – Boolean (Heyting) algebra, including representation theorems for distributive lattices and Heyting algebras. We regard this Chapter as preparatory to Chapters 10 and 12 in which rough sets are discussed also in an abstract setting of lattices.

The last in this part Chapter 9 concerns predicate calculus. We give a naive introduction as well as a formalization of the predicate calculus both in the Rasiowa – Sikorski style and in the Gentzen style. In addition, intuitionistic logic is introduced as a preliminary to Chapter 10, and the Łukasiewicz calculus of fractional values in logic is recalled to be used in Chapter 10. As a preliminary to Chapter 12, Lindenbaum –Tarski algebras of logical calculi are introduced and studied in some detail.

After preliminaries are set out, we proceed with an advanced discussion of mathematical results pertaining to rough sets. In Chapter 10, we discuss independence issues, covering the Pawlak – Novotný theory of reducts and independence via semi – lattices. After that we discuss an abstract theory of approximations and the theory of partial dependencies referring to the Łukasiewicz calculus of fractional values (Chapter 9).

Chapter 11 is devoted to topological theory of rough sets. In the setting of information systems with countably many attributes, we construct topologies on rough sets and almost rough sets metrizable by metrics D, D^*, D' making rough sets into complete metric spaces. As an application, we introduce the approximate fixed point theorem on rough set spaces and we discuss for the first time in literature the notion of a fractal in information systems showing continuity of fractal dimension with respect to induced metrics.

In Chapter 12, a study of algebraic properties of rough sets is presented. We begin with the Lindenbaum – Tarski algebra of the 3–valued Łukasiewicz logic i.e. with the Wajsberg algebra and we introduce Łukasiewicz algebras. We include a proof of equivalence of Wajsberg algebras to 3 – Łukasiewicz algebras. Then we demonstrate that rough sets in disjoint representation may be given a structure of a Łukasiewicz algebra i.e. that they are related in a natural way to the 3 – valued logic of Łukasiewicz. Parallel representations of rough sets as a Heyting algebra, a double regular Stone algebra or a Post algebra due to Pagliani, Pomykala et al. are also introduced. After that we pursue the logical thread and we present the Rauszer results about the equivalence between calculus of independence and the fragment of intuitionistic logic. We conclude with an account of modalities in information systems presenting the information logic *IL* due to Vakarelov.

At this point our presentation of rough set mathematics *sensu stricto* ends. We come back to the starting idea in this preface viz. to relations between rough and fuzzy set theories.

In order to give the reader an insight into basic differences in technical approaches to either theory, we present in Chapter 13 an account of infinite valued logic, based on an axiomatics due to Łukasiewicz. We present a syntactic proof of its completeness after Rose–Rosser with some simplifications concerning matrix dichotomies which we replace with some simple algebraic lemmas.

After that we present an account of a fuzzy sentential logic due to Pavelka in which we introduce adjoint pairs being a logical counterpart of pairs (t – norm, residuated implication) so essential in fuzzy set calculi. Working with the Łukasiewicz adjoint pair, we include the proof of completeness of the fuzzy sentential logic in the Pavelka sense.

Chapter 13 introduces the reader to mathematical world associated with the fuzzy set theory and contrasted with the 3–valued world of rough set theory.

This contrast is underlined in Chapter 14 where the ideas of rough fuzzy set and fuzzy rough set are introduced. We begin with a discussion of t – norms and t – conorms along with induced residuated implications and we prove the representation theorem about t – norms and t – conorms from which we infer the remarkable fact that the Łukasiewicz adjoint pair is the unique up to equivalence adjoint pair making the fuzzy sentential logic complete.

Then we introduce rough fuzzy sets in the Dubois – Prade style followed by fuzzy rough sets. Here the notions of a fuzzy equivalence relation as well as a fuzzy partition are discussed after Zadeh et al., and a proposal for fuzzy rough sets is laid out. Finally, we return to the lattice setting and we discuss the Brouwer – Zadeh lattices of fuzzy sets with rough set style approximations equivalent to the certainty and possibility regions of a fuzzy sets mentioned in the introduction to this Preface.

It may be noticed from the outline of the content presented above that our book covers all essential results in mathematical theory of rough sets which require a deeper mathematical knowledge.

The book may be used as a text for a one – semester graduate course on rough set theory, exploiting Chapters 10 – 12 with a possible addition of excerpts from Chapters 13, 14.

Due to Chapters 2-9 the book may be also used as a text for one – semester undergraduate course covering less demanding topics of Chapters 10 – 12 along with fragments of Chapters 2–9 as a prerequisite.

Finally the book may be used as a text for a course on mathematical foundations of Computer Science as it covers the basic facts from logic, set theory, topology and lattice theory.

Over 320 exercises in Chapters 2-13 provided with careful hints allow the reader to gain additional information. The basic text along with exercises makes the book completely self – contained and the diligent reader will master the topic on its own without the need to refer to other sources if they are pleased to do so.

One more usage of the book will be as a reference to the area of mathematical foundations of rough set theory as each Chapter is provided with the essential bibliography. The author hopes that the book will provide many newcomers to rough set theory with the necessary background on the current state of art in this area and by this it will stimulate the further development of the theory.

The author has experienced the friendship and help of many during his work on rough sets.

The thanks should go first to the late Professor Helena Rasiowa, an eminent logician, a disciple of Jan Łukasiewicz whose legacy so permeates this book, who brought the author to the rough set theory, and to Professor Zdzisław Pawlak, the founder of the rough set theory for His kind help and support. This book is offered to him.

The author would like to thank Professor Andrzej Skowron for his help and support, and for many years of close and still going into future collaboration on rough sets which has made those years so pleasant.

Professor Janusz Kacprzyk has always supported the author in his editorial tasks which is acknowledged with gratitude.

Colleagues from the Polish–Japanese Institute of Information Technology has provided a friendly atmosphere and facilities which allowed for preparing of this book.

I am also thinking of my wife Maria and my son Martin in grateful acknowledgment of their support and love.

Finally the author thanks Dr Martina Bihn and her colleagues from Physica–Verlag/Springer–Verlag for their skilled help in making this book.

This work was supported by the Grant from the Polish State Committee for Scientific Research (KBN) No. 8T11C 024 17. A help from the Grant Nr 8T11C 02519 from the Polish State Committee for Scientific Research(KBN) is also acknowledged.

For the reader convenience a guide map of the content with suggestions for connections among Chapters is offered. A graduate course could cover 1-10-11-12-13-14, a less demanding undergraduate course could encompass the same without more technical proofs. A course on Mathematics for Computer Science could cover 2-3-4-5-6-7-8-9 or 2-3-4-5-9. A course on Fuzzy Set Theory devoted to foundational issues may be organized around Chapter 13 and Chapter 14.

The author does hope that this book will offer the rough set community a useful reference book and researchers from other areas of investigation, in particular from the field of fuzzy set theory, will find therein a basic information on rough set theory, its purposes, methods, and mathematical tools it does require as well as on relations between the two theories.

April, 2002 Lech Polkowski

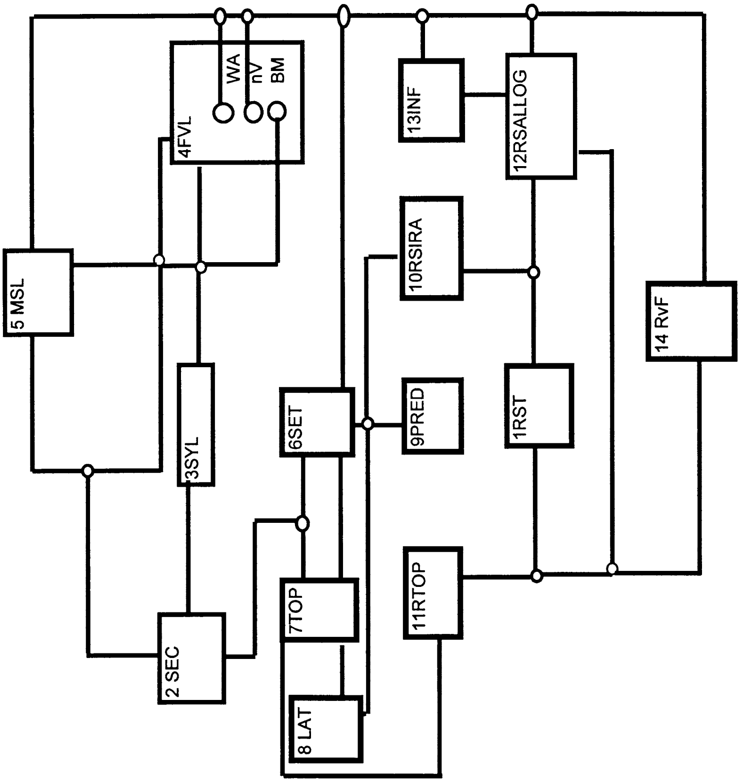

Figure 1: A guide map: WA is Sections 4.1-4.4, nV is Sections 4.5-6, BM is Section 4.7; 14RvF is Chapter 14

Contents