LNCS 4100 Journal Subline

Transactions on
Rough Sets V

James F. Peters · Andrzej Skowron

Editors-in-Chief

Springer

James F. Peters Andrzej Skowron (Eds.)

Transactions on Rough Sets V

Springer

Volume Editors

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

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

Library of Congress Control Number: 2006933430

CR Subject Classification (1998): F.4.1, F.1, I.2, H.2.8, I.5.1, I.4

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

ISSN 1861-2059
ISBN-10 3-540-39382-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-39382-5 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11847465 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 4100

Preface

Volume V of the Transactions on Rough Sets (TRS) is dedicated to the monumental life and work of Zdzisław Pawlak[1]. During the past 35 years, since the introduction of knowledge description systems in the 1970s, the theory and applications of rough sets have grown rapidly. This volume continues the tradition begun with earlier volumes of the TRS series and introduces a number of new advances in the foundations and application of rough sets. These advances have profound implications in a number of research areas such as adaptive learning, approximate reasoning and belief systems, approximation spaces, Boolean reasoning, classification methods, classifiers, concept analysis, data mining, decision logic, decision rule importance measures, digital image processing, recognition of emotionally-charged gestures in animations, flow graphs, Kansei engineering, movie sound track restoration, multicriteria decision analysis, relational information systems, rough-fuzzy sets, rough measures, signal processing, variable precision rough set model, and video retrieval. It can be observed from the papers included in this volume that research concerning the foundations and applications of rough sets remains an intensely active research area worldwide. A total of 37 researchers from 8 countries are represented in this volume, the countries being, Canada, India, P.R. China, Poland, Japan, Taiwan, UK and the USA.

A capsule view of the life and work of Zdzisław Pawlak is included in an article at the beginning of this volume. During his lifetime, the research interests of Pawlak were rich and varied. His research ranged from his pioneering work on knowledge description systems and rough sets during the 1970s and 1980s as well as his work on the design of computers, information retrieval, modeling conflict analysis and negotiation, genetic grammars and molecular computing. Added to that was Pawlak's lifelong interest in painting, photography and poetry. During his lifetime, Pawlak nurtured worldwide interest in approximation, approximate reasoning and rough set theory and its applications. Evidence of the influence of Pawlak's work can been seen in the growth in the rough-set literature that now includes over 4000 publications as well as the growth and maturity of the International Rough Set Society.

TRS V also includes 15 papers that explore the theory of rough sets as well as new applications of rough sets. In addition, this volume of the TRS includes a complete monograph on rough sets and approximate Boolean reasoning systems that includes both the foundations as well as the applications of data mining, by Hung Son Nguyen. New developments in the foundations of rough sets are represented by a number of papers in this volume, namely, Rough Truth, Consistency and Belief Change (Mohua Banerjee), Rough Set Approximations in Formal Concept Analysis (Yiyu Yao and Yaohua Chen), Rule Importance Measures (Jiye Li and Nick Cercone), Generalized Rough-Fuzzy Approximation Operators (Wei-Zhi Wu), Rough Set Flow Graphs (Cory Butz, W. Yan, and

[1] Prof. Pawlak passed away on 7 April 2006.

B. Yang), Vague Concept Approximation and Adaptive Learning (Jan Bazan, Andrzej Skowron, and Roman Świniarski), and Arrow Decision Logic (Tuan-Fang Fan, Duen-Ren Liu, and Gwo-Hshiung Tzeng). Applications of rough sets are also represented by the following papers in this volume: Matching 2D Image Segments with Genetic Algorithms and Approximations Spaces (Maciej Borkowski and James Peters), Rough Set-Based Application to Recognition of Emotionally-Charged Animated Characters Gestures (Bożena Kostek and Piotr Szczuko), Movie Sound Track Restoration (Andrzej Czyżewski, Marek Dziubinski, Lukasz Litwic, and Przemyslaw Maziewski), Multimodal Classification Case Studies (Andrzej Skowron, Hui Wang, Arkadiusz Wojna, and Jan Bazan), P300 Wave Detection Using Rough Sets (Sheela Ramanna and Reza Fazel Rezai), Motion-Information-Based Video Retrieval Using Rough Pre-classification (Zhe Yuan, Yu Wu, Guoyin Wang, and Jianbo Li), Variable Precision Baysian Rough Set Model and Its Application to Kansei Engineering (Tatsuo Nishino, Mitsuo Nagamachi, and Hideo Tanaka).

The Editors of this volume extend their hearty thanks to the reviewers of the papers that were submitted to this TRS volume: Mohua Banerjee, Jan Bazan, Teresa Beauboeuf, Maciej Borkowski, Gianpiero Cattaneo, Nick Cercone, Davide Cuicci, Andrzej Czyżewski, Jitender Deogun, Ivo Düntsch, Reza Fazel-Rezai, Anna Gomolińska, Jerzy Grzymała-Busse, Masahiro Iniguichi, Jouni Järvinen, Mieczysław Kłopotek, Beata Konikowska, Bożena Kostek, Marzena Kryszkiewicz, Rafał Latkowski, Churn-Jung Liau, Pawan Lingras, Jan Małuszyński, Benedetto Matarazzo, Michał Mikołajczyk, Mikhail Moshkov, Maria Nicoletti, Hoa Nguyen, Son Nguyen, Piero Pagliani, Sankar Pal, Witold Pedrycz, Lech Polkowski, Anna Radzikowska, Vijay Raghavan, Sheela Ramanna, Zbigniew Raś, Dominik Ślęzak, Jerzy Stefanowski, Jarosław Stepaniuk, Zbigniew Suraj, Roman Świniarski, Piotr Synak, Marcin Szczuka, Daniel Vanderpooten, Dimiter Vakarelov, Alicja Wieczorkowska, Arkadiuz Wojna, Marcin Wolski, Jakub Wróblewski, Dan Wu, Wei-Zhi Wu, Yiyu Yao, and Wojciech Ziarko.

This issue of the TRS was made possible thanks to the reviewers as well as to the laudable efforts of a great many generous persons and organizations. The editors and authors of this volume also extend an expression of gratitude to Alfred Hofmann, Ursula Barth, Christine Günther and the other LNCS staff at Springer for their support in making this volume of the TRS possible. In addition, the editors of this volume extend their thanks to Dominik Ślęzak for his help and suggestions concerning extensions of selected RSFDGrC 2005 papers included in this volume of the TRS. We anticipate that additional RSFDGrC 2005 papers now being reviewed will be included in future volumes of the TRS. We also extend our thanks to Marcin Szczuka for his consummate skill and care in the compilation of this volume. The Editors of this volume have been supported by the Ministry for Scientific Research and Information Technology of the Republic of Poland, research grant No. 3T11C00226, and the Natural Sciences and Engineering Research Council of Canada (NSERC) research grant 185986 respectively.

June 2006

James F. Peters
Andrzej Skowron

LNCS Transactions on Rough Sets

This journal subline has as its principal aim the fostering of professional exchanges between scientists and practitioners who are interested in the foundations and applications of rough sets. Topics include foundations and applications of rough sets as well as foundations and applications of hybrid methods combining rough sets with other approaches important for the development of intelligent systems.

The journal includes high-quality research articles accepted for publication on the basis of thorough peer reviews. Dissertations and monographs up to 250 pages that include new research results can also be considered as regular papers. Extended and revised versions of selected papers from conferences can also be included in regular or special issues of the journal.

Lecture Notes in Computer Science

For information about Vols. 1–4143

please contact your bookseller or Springer

Vol. 4190: R. Larsen, M. Nielsen, J. Sporring (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006, Part I. XXXVVIII, 949 pages. 2006.

Vol. 4189: D. Gollmann, J. Meier, A. Sabelfeld (Eds.), Computer Security – ESORICS 2006. XI, 548 pages. 2006.

Vol. 4188: P. Sojka, I. Kopeček, K. Pala (Eds.), Text, Speech and Dialogue. XIV, 721 pages. 2006. (Sublibrary LNAI).

Vol. 4187: J.J. Alferes, J. Bailey, W. May, U. Schwertel (Eds.), Principles and Practice of Semantic Web Reasoning. XI, 277 pages. 2006.

Vol. 4186: C. Jesshope, C. Egan (Eds.), Advances in Computer Systems Architecture. XIV, 605 pages. 2006.

Vol. 4185: R. Mizoguchi, Z. Shi, F. Giunchiglia (Eds.), The Semantic Web – ASWC 2006. XX, 778 pages. 2006.

Vol. 4184: M. Bravetti, M. Núñez, G. Zavattaro (Eds.), Web Services and Formal Methods. X, 289 pages. 2006.

Vol. 4183: J. Euzenat, J. Domingue (Eds.), Artificial Intelligence: Methodology, Systems, and Applications. XIII, 291 pages. 2006. (Sublibrary LNAI).

Vol. 4182: H.T. Ng, M.-K. Leong, M.-Y. Kan, D. Ji (Eds.), Information Retrieval Technology. XVI, 684 pages. 2006.

Vol. 4180: M. Kohlhase, OMDoc – An Open Markup Format for Mathematical Documents [version 1.2]. XIX, 428 pages. 2006. (Sublibrary LNAI).

Vol. 4179: J. Blanc-Talon, W. Philips, D. Popescu, P. Scheunders (Eds.), Advanced Concepts for Intelligent Vision Systems. XXIV, 1224 pages. 2006.

Vol. 4178: A. Corradini, H. Ehrig, U. Montanari, L. Ribeiro, G. Rozenberg (Eds.), Graph Transformations. XII, 473 pages. 2006.

Vol. 4177: R. Marín, E. Onaindía, A. Bugarín, J. Santos (Eds.), Current Topics in Artificial Intelligence. XIII, 621 pages. 2006. (Sublibrary LNAI).

Vol. 4176: S.K. Katsikas, J. Lopez, M. Backes, S. Gritzalis, B. Preneel (Eds.), Information Security. XIV, 548 pages. 2006.

Vol. 4175: P. Bücher, B.M.E. Moret (Eds.), Algorithms in Bioinformatics. XII, 402 pages. 2006. (Sublibrary LNBI).

Vol. 4174: K. Franke, K.-R. Müller, B. Nickolay, R. Schäfer (Eds.), Pattern Recognition. XX, 773 pages. 2006.

Vol. 4173: S. El Yacoubi, B. Chopard, S. Bandini (Eds.), Cellular Automata. XV, 734 pages. 2006.

Vol. 4172: J. Gonzalo, C. Thanos, M. F. Verdejo, R.C. Carrasco (Eds.), Research and Advanced Technology for Digital Libraries. XVII, 569 pages. 2006.

Vol. 4169: H.L. Bodlaender, M.A. Langston (Eds.), Parameterized and Exact Computation. XI, 279 pages. 2006.

Vol. 4168: Y. Azar, T. Erlebach (Eds.), Algorithms – ESA 2006. XVIII, 843 pages. 2006.

Vol. 4167: S. Dolev (Ed.), Distributed Computing. XV, 576 pages. 2006.

Vol. 4166: J. Górski (Ed.), Computer Safety, Reliability, and Security. XIV, 440 pages. 2006.

Vol. 4165: W. Jonker, M. Petković (Eds.), Secure, Data Management. X, 185 pages. 2006.

Vol. 4163: H. Bersini, J. Carneiro (Eds.), Artificial Immune Systems. XII, 460 pages. 2006.

Vol. 4162: R. Královič, P. Urzyczyn (Eds.), Mathematical Foundations of Computer Science 2006. XV, 814 pages. 2006.

Vol. 4161: R. Harper, M. Rauterberg, M. Combetto (Eds.), Entertainment Computing - ICEC 2006. XXVII, 417 pages. 2006.

Vol. 4160: M. Fisher, W.v.d. Hoek, B. Konev, A. Lisitsa (Eds.), Logics in Artificial Intelligence. XII, 516 pages. 2006. (Sublibrary LNAI).

Vol. 4159: J. Ma, H. Jin, L.T. Yang, J.J.-P. Tsai (Eds.), Ubiquitous Intelligence and Computing. XXII, 1190 pages. 2006.

Vol. 4158: L.T. Yang, H. Jin, J. Ma, T. Ungerer (Eds.), Autonomic and Trusted Computing. XIV, 613 pages. 2006.

Vol. 4156: S. Amer-Yahia, Z. Bellahsène, E. Hunt, R. Unland, J.X. Yu (Eds.), Database and XML Technologies. IX, 123 pages. 2006.

Vol. 4155: O. Stock, M. Schaerf (Eds.), Reasoning, Action and Interaction in AI Theories and Systems. XVIII, 343 pages. 2006. (Sublibrary LNAI).

Vol. 4154: Y.A. Dimitriadis, I. Zigurs, E. Gómez-Sánchez (Eds.), Groupware: Design, Implementation, and Use. XIV, 438 pages. 2006.

Vol. 4153: N. Zheng, X. Jiang, X. Lan (Eds.), Advances in Machine Vision, Image Processing, and Pattern Analysis. XIII, 506 pages. 2006.

Vol. 4152: Y. Manolopoulos, J. Pokorný, T. Sellis (Eds.), Advances in Databases and Information Systems. XV, 448 pages. 2006.

Vol. 4151: A. Iglesias, N. Takayama (Eds.), Mathematical Software - ICMS 2006. XVII, 452 pages. 2006.

Vol. 4150: M. Dorigo, L.M. Gambardella, M. Birattari, A. Martinoli, R. Poli, T. Stützle (Eds.), Ant Colony Optimization and Swarm Intelligence. XVI, 526 pages. 2006.

Vol. 4149: M. Klusch, M. Rovatsos, T.R. Payne (Eds.), Cooperative Information Agents X. XII, 477 pages. 2006. (Sublibrary LNAI).

Vol. 4148: J. Vounckx, N. Azemard, P. Maurine (Eds.), Integrated Circuit and System Design. XVI, 677 pages. 2006.

Vol. 4147: M. Broy, I.H. Krüger, M. Meisinger (Eds.), Automotive Software – Connected Services in Mobile Networks. XIV, 155 pages. 2006.

Vol. 4146: J.C. Rajapakse, L. Wong, R. Acharya (Eds.), Pattern Recognition in Bioinformatics. XIV, 186 pages. 2006. (Sublibrary LNBI).

Vol. 4145: L. Moreau, I. Foster (Eds.), Provenance and Annotation of Data and Processes. XI, 288 pages. 2006.

Vol. 4144: T. Ball, R.B. Jones (Eds.), Computer Aided Verification. XV, 564 pages. 2006.

¥547.00元

Table of Contents

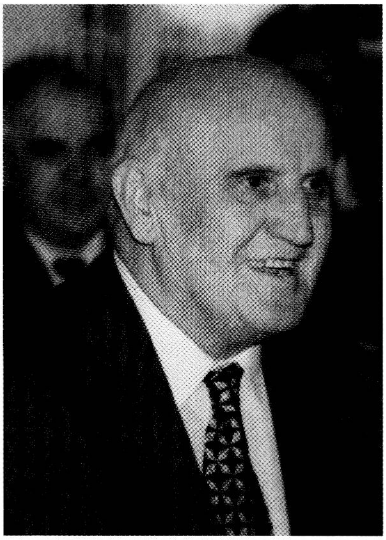
Commemoration

Regular Papers

Dissertations and Monographs

Zdzisław Pawlak: Life and Work
1926-2006

James F. Peters and Andrzej Skowron



In the history of mankind, Professor Zdzisław Pawlak, Member of the Polish Academy of Sciences, will be remembered as a great human being with exceptional humility, wit and kindness as well as an extraordinarily innovative researcher with exceptional stature. His legacy is rich and varied. Pawlak's research contributions have had far-reaching implications inasmuch as his works are fundamental in establishing new perspectives for scientific research in a wide spectrum of fields.

Preamble

Professor Pawlak's most widely recognized contribution is his incisive approach to classifying objects with their attributes (features) and his introduction of approximation spaces, which establish the foundations of granular computing and provide frameworks for perception and knowledge discovery in many areas. He was with us only for a short time and, yet, when we look back at his accomplishments, we realize how greatly he has influenced us with his generous spirit and creative work in many areas such as approximate reasoning, intelligent systems research, computing models, mathematics (especially, rough set theory), molecular computing, pattern recognition, philosophy, art, and poetry. This article attempts to give a vignette that highlights some of Pawlak's remarkable accomplishments. This vignette is limited to a brief coverage of Pawlak's work in rough set theory, molecular computing, philosophy, painting and poetry. Detailed coverage of these as well as other accomplishments by Pawlak is outside the scope of this commemorative article.

1 Introduction

This article commemorates the life, work and creative genius of Zdzisław Pawlak. He is well-known for his innovative work on the classification of objects by means

J.F. Peters and A. Skowron (Eds.): Transactions on Rough Sets V, LNCS 4100, pp. 1–24, 2006.

of attributes (features) [25] and his discovery of rough set theory during the early 1980s (see, e.g., [11,22,25,27]). Since the introduction of rough set theory, there have been well over 4000 publications on this theory and its applications (see, e.g., [6,35,36,37,39,71] and Section 12).

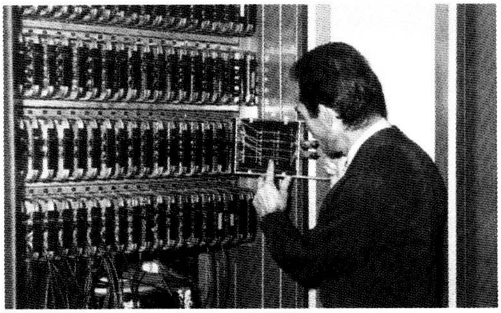
One can also observe a number of other facets of Pawlak's life and work that are less known, namely, his pioneering work on genetic grammars and molecular computing, his interest in philosophy, his lifelong devotion to painting landscapes and waterscapes depicting the places he visited, his interest and skill in photography, and his more recent interests in poetry and methods of solving mysteries by fictional characters such as Sherlock Holmes. During his life, Pawlak contributed to the foundations of granular computing, intelligent systems research, computing models, mathematics (especially, rough set theory), molecular computing, knowledge discovery as well as knowledge representation, and pattern recognition.

This article attempts to give a brief vignette that highlights some of Pawlak's remarkable accomplishments. This vignette is limited to a brief coverage of Pawlak's works in rough set theory, molecular computing, philosophy, painting and poetry. Detailed coverage of these as well as other accomplishments by Pawlak is outside the scope of this commemorative article.

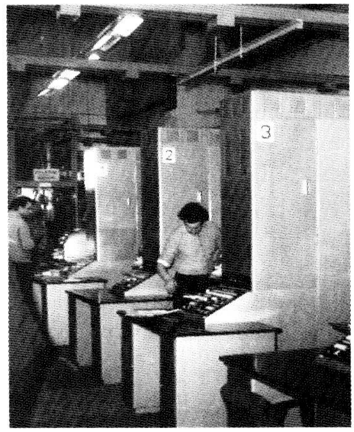
The article is organized as follows. A brief biography of Zdzisław Pawlak is given in Sect. 2. Some of the very basic ideas of Pawlak's rough set theory are presented in Sect. 3. This is followed by a brief presentation of Pawlak's introduction of a genetic grammar and molecular computing in Sect. 8. Pawlak's more recent reflections concerning philosophy (especially, the philosophy of mathematics) are briefly covered in Sect. 9. Reflections on Pawlak's lifelong interest in painting and nature as well as a sample of paintings by Pawlak and a poem coauthored by Pawlak, are presented in Sect. 10.

2 Zdzisław Pawlak: A Brief Biography

Zdzisław Pawlak was born on 10 November 1926 in Łódź, 130 km south-west from Warsaw, Poland [41]. In 1947, Pawlak began studying in the Faculty of Electrical Engineering at Łódź University of Technology, and in 1949 continued his studies in the Telecommunication Faculty at Warsaw University of Technology. Starting in the early 1950s and continuing throughout his life, Pawlak painted the places he visited, especially landscapes and waterscapes reflecting his observations in Poland and other parts of the world. This can be seen as a continuation of the work of his father, who was fond of wood carving and who carved a wooden self-portrait that was kept in Pawlak's study. He also had extraordinary skill in mathematical modeling in the organization of systems (see, e.g., [20,24,28]) and in computer systems engineering (see, e.g., [16,17,18,19,21]). During his early years, he was a pioneer in the designing computing machines. In 1950, Pawlak presented the first project of a computer called GAM 1. He completed his M.Sc. in Telecommunication Engineering in 1951. Pawlak's publication in 1956 on a new method for random number generation was the first article in informatics



1.1: Interior of UMC1



1.2: UMC1 Prototype

Fig. 1. Snapshots of the UMC1 Computer System

published abroad by a researcher from Poland [13]. In 1958, Pawlak completed his doctoral degree from the Institute of Fundamental Technological Research at the Polish Academy of Science with a Thesis on *Applications of Graph Theory to Decoder Synthesis.* In 1961, Pawlak was also a member of a research team that constructed one of the first computers in Poland called UMC 1 (see Fig. 1).

The original arithmetic for the UMC1 computer system with base "-2" was due to Pawlak [14]. He received his habilitation from the Institute of Mathematics at the Polish Academy of Sciences in 1963. In his habilitation entitled *Organization of Address-Less Machines*, Pawlak proposed and investigated parenthesis-free languages, a generalization of polish notation introduced by Jan Łukasiewicz (see, e.g., [16,17]).

In succeeding years, Pawlak worked at the Institute of Mathematics of Warsaw University and, in 1965, introduced foundations for modeling DNA [15] in what has come to be known as molecular computing [3,15]. He also proposed a new formal model of a computing machine known as the Pawlak machine [21,23] that is different from the Turing machine and from the von Neumann machine. In 1973, he introduced knowledge representation systems [22] as part of his work on the mathematical foundations of information retrieval (see, e.g., [11,22]). In the early 1980s, he was part of a research group at the Institute of Computer Science of the Polish Academy of Sciences, where he discovered rough sets and the idea of classifying objects by means of their attributes [25], which was the basis for extensive research in rough set theory during the 1980s (see, e.g., [7,8,12,26,27, 29]). During the succeeding years, Pawlak refined and amplified the foundations of rough sets and their applications, and nurtured worldwide research in rough sets that has led to over 4000 publications (see, e.g., [39]). In addition, he did extensive work on the mathematical foundations of information systems during the early 1980s (see, e.g., [24,28]). He also invented a new approach to conflict analysis (see, e.g., [30,31,33,34]).

During his later years, Pawlak's interests were very diverse. He developed a keen interest in philosophy, especially in the works by Łukasiewicz (logic and probability), Leibniz (*identify of indiscernibles*), Frege (membership, sets), Russell (antinomies), and Leśniewski (*being a part*)). Pawlak was also interested in the works of detective fiction by Sir Arthur Conan Doyle (especially, Sherlock Holmes' fascination with data as a basis for solving mysteries) (see, e.g., [35]).

Finally, Zdzisław Pawlak gave generously of his time and energy to help others. His spirit and insights have influenced many researchers worldwide. During his life, he manifested an extraordinary talent for inspiring his students and colleagues as well as many others outside his immediate circle. For this reason, he was affectionately known to some of us as Papa Pawlak.

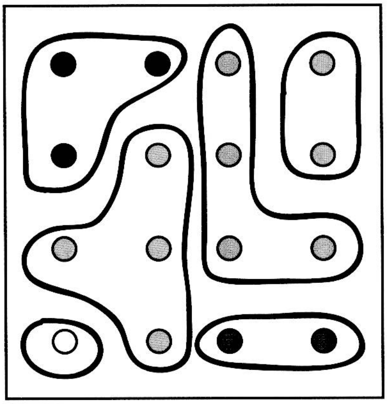
3 Rough Sets

> If we classify objects by means of attributes,
> exact classification is often impossible.
> – Zdzisław Pawlak, January 1981.

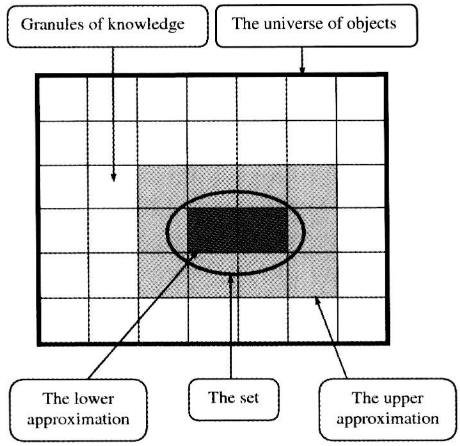
A brief presentation of the foundations of rough set theory is given in this section. Rough set theory has its roots in Zdzisław Pawlak's research on knowledge representation systems during the early 1970s [22]. Rather than attempt to classify objects *exactly* by means of attributes (features), Pawlak considered an approach to solving the object classification problem in a number of novel ways. First, in 1973, he formulated knowledge representation systems (see, e.g., [11,22]). Then, in 1981, Pawlak introduced approximate descriptions of objects and considered knowledge representation systems in the context of upper and lower classification of objects relative to their attribute values [25,26]. We start with a system $S = (X, A, V, \sigma)$, where X is a non-empty set of objects, A is a set of attributes, V is a union of sets V_a of values associated with each $a \in A$, and σ is called a knowledge function defined as the mapping $\sigma : X \times A \rightarrow V$, where $\sigma(x, a) \in V_a$ for every $x \in X$ and $a \in A$. The function σ is referred to as *knowledge function* about objects from X. The set X is partitioned into elementary sets that later were called blocks, where each elementary set contains those elements of X which have matching attribute values. In effect, a block (elementary set) represents a granule of knowledge (see Fig. 2.2). For example, for any $B \subseteq A$ the B-elementary set for an element $x \in X$ is denoted by $B(x)$, which is defined by

$$B(x) = \{y \in X |\ \forall a \in B\ \sigma(x, a) = \sigma(y, a)\} \quad (1)$$

Consider, for example, Fig. 2.1 which represents a system S containing a set X of colored circles and a feature set A that contains only one attribute, namely, *color*. Assume that each circle in X has only one color. Then the set X is partitioned into elementary sets or blocks, where each block contains circles with the same color. In effect, elements of a set $B(x) \subseteq X$ in a system S are classified as *indiscernible* if they are indistinguishable by means of their feature values for any $a \in B$. A set of *indiscernible* elements is called an *elementary set* [25]. Hence,



2.1: Blocks of Objects



2.2: Sample Set Approximation

Fig. 2. Rudiments of Rough Sets

any subset $B \subseteq A$ determines a partition $\{B(x) : x \in X\}$ of X. This partition defines an equivalence relation $Ind(B)$ on X called an *indiscernibility* relation such that $xInd(B)y$ if and only if $y \in B(x)$ for every $x, y \in X$. Assume that $Y \subseteq X$ and $B \subseteq A$, and consider an approximation of the set Y by means of the attributes in B and B-indiscernible blocks in the partition of X. The union of all blocks that constitute a subset of Y is called the *lower approximation* of Y (usually denoted by B_*Y), representing certain knowledge about Y. The union of all blocks that have non-empty intersection with the set Y is called the *upper approximation* of Y (usually denoted by B^*Y), representing uncertain knowledge about Y. The set $BN_B(Y) = B^*Y - B_*Y$ is called the B-boundary of the set Y. In the case where $BN_B(Y)$ is non-empty, the set Y is a *rough (imprecise)* set. Otherwise, the set Y is a *crisp* set. This approach to classification of objects in a set is represented graphically in Fig. 2.2, where the region bounded by the ellipse represents a set Y, the darkened blocks inside Y represent B_*Y, the gray blocks represent the boundary region $BN_B(Y)$, and the gray and the darkened blocks taken together represent B^*Y.

Consequences of this approach to the classification of objects by means of their feature values have been remarkable and far-reaching. Detailed accounts of the current research in rough set theory and its applications are available, e.g., in [35, 36, 37] (see also Section 12).

4 Approximation

> Some categories (subsets of objects) cannot be
> expressed exactly by employing available knowledge.
> Hence, we arrive at the idea of approximation
> of a set by other sets.
>
> –Zdzisław Pawlak, 1991.

One of the most profound, very important notions underlying rough set theory is approximation. In general, an *approximation* is defined as the replacement of objects by others that resemble the original objects in certain respects [4]. For example, consider a universe U containing objects representing behaviors of agents. In that case, we can consider blocks of behaviors in the partition U/R, where the behaviors within a block resemble (are *indiscernible* from) each other by virtue of their feature values. Then any subset X of U can be approximated by blocks that are either proper subsets of X (lower approximation of the set X denoted $\underline{R}X$) or by blocks having one or more elements in common with X (upper approximation of the set X denoted $\overline{R}X$)[1]. In rough set theory, the focus is on approximating one set of objects by means of another set of objects based on the feature values of the objects [32]. The lower approximation operator $\underline{R}$ has properties that correspond closely to properties of what is known as the Π_0 topological *interior* operator [27, 77]. Similarly, the upper approximation operator $\overline{R}$ has properties that correspond closely to properties of the Π_0 topological *closure* operator [27, 77]. It was observed in [27] that the key to the rough set approach is provided by the exact mathematical formulation of the concept of approximative (rough) equality of sets in a given approximation space.

5 Approximation Spaces

> The key to the presented approach
> is provided by the exact mathematical formulation,
> of the concept of approximative (rough) equality
> of sets in a given approximation space.
> –Zdzisław Pawlak, 1982.

In [32], an approximation space is represented by the pair (U, R), where U is a universe of objects, and $R \subseteq U \times U$ is an indiscernibility relation (denoted *Ind* as in Sect. 3) defined by an attribute set (i.e., $R = Ind(A)$ for some attribute set A). In this case, R is an equivalence relation. Let $[x]_R$ denote an equivalence class of an element $x \in U$ under the indiscernibility relation R, where $[x]_R = \{y \in U : xRy\}$.

In this context, R-approximations of any set $X \subseteq U$ are based on the exact (crisp) containment of sets. Then set approximations are defined as follows:

- $x \in U$ belongs with certainty to $X \subseteq U$ (i.e., x belongs to the R-lower approximation of X), if $[x]_R \subseteq X$.
- $x \in U$ possibly belongs $X \subseteq U$ (i.e., x belongs to the R-upper approximation of X), if $[x]_R \cap X \neq \emptyset$.

[1] In more recent years, the notation R_*X, R^*X has been often used (see, e.g., Sect. 3) to denote lower and upper approximation, respectively, since this notation is more "typewriter" friendly.