KNOWLEDGE-BASED CLUSTERING

From Data to Information Granules

Witold Pedrycz

Department of Electrical and Computer Engineering
University of Alberta
Edmonton, Canada

and

Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

A JOHN WILEY & SONS, INC., PUBLICATION

KNOWLEDGE-BASED CLUSTERING

KNOWLEDGE-BASED CLUSTERING

From Data to Information Granules

Witold Pedrycz

Department of Electrical and Computer Engineering
University of Alberta
Edmonton, Canada

and

Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Pedrycz, Witold, 1953-
Knowledge-based clustering : from data to information granules / Witold Pedrycz.
p. cm.
"A Wiley-Interscience publication."
Includes bibliographical references and index.
ISBN 0-471-46966-1 (cloth)
1. Soft computing. 2. Granular computing. 3. Fuzzy systems. I. Title.
QA76.9.S63P45 2005
006.3–dc22

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To Ewa, Adam, and Barbara

Contents

Foreword

It is always a challenging task to write a foreword to a work authored by Professor Pedrycz. The reason is that, as a rule, what he writes about goes far beyond what can be found in the existing literature. This is particularly true in the instance of *Knowledge-Based Clustering: From Data to Information Granules* or *Knowledge-Based Clustering*, for short. *Knowledge-Based Clustering* is a magnum opus which touches upon some of the most basic facets of human cognition. It does so with authority, originality, erudition, insight, and high expository skill. Profusion of examples, figures, and references make Professor Pedrycz's work a pleasure to read.

In *Knowledge-Based Clustering*, Professor Pedrycz addresses a vast array of linked subjects. Starting with an exposition of clustering and fuzzy clusters, he moves to computing with granular information, with a granule being a clump of attribute-values drawn together by indistinguishability, equivalence, similarity, proximity, or functionality. Professor Pedrycz's co-authorship of a recent text on granular computing provides him with an effective framework for linking clustering with granular computing. In granular computing, the objects of computation are granules rather than singletons. In its general form, granular computing subsumes computing with intervals, computing with rough sets, and computing with probability distributions. The linkage between granular computing and cluster analysis plays a pivotal role throughout Professor Pedrycz's work, and is an important novel feature of his approach to cluster analysis.

The chapters that are focused on granular computing serve as a foundation for the core of the book—knowledge-based clustering. In this mode of clustering, clustering is guided by the knowledge that underlies data. There is much that is new in this part of the book, especially in chapters dealing with conditioned fuzzy clustering, collaborative clustering, directional clustering, fuzzy relational clustering, and clustering of nonhomogeneous patterns. The last part of Professor Pedrycz's work is an informative exposition of applications of knowledge-based clustering to generic models. In this part, we find a range of unconventional concepts and techniques, among them hyperbox modeling, linguistic modeling, and granular mapping.

To see the importance of Pedrycz's work in a proper perspective, an observation is in order. As we move further into the age of machine intelligence and automated reasoning, a daunting problem becomes harder and harder to master. How can we cope with the explosive growth in data, information, and knowledge? How can we locate and infer from decision-relevant information that is embedded in a large database that is unstructured, imprecise, and not totally reliable?