Vito Di Gesù
Francesco Masulli
Alfredo Petrosino (Eds.)

LNAI 2955

Fuzzy Logic and Applications

5th International Workshop, WILF 2003
Naples, Italy, October 2003
Revised Selected Papers

Springer

Vito Di Gesù Francesco Masulli
Alfredo Petrosino (Eds.)

Fuzzy Logic and Applications

5th International Workshop, WILF 2003
Naples, Italy, October 9-11, 2003
Revised Selected Papers

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Vito Di Gesù
Università degli Studi di Palermo
Dipartimento di Matematica ed Applicazioni
Via Archirafi, 34, 90123 Palermo, Italy
E-mail: digesu@math.unipa.it

Francesco Masulli
Polo Universitario di La Spezia, DISI
Via Dodecaneso 35, 16146 Genoa, Italy
E-mail: masulli@disi.unige.it

Alfredo Petrosino
University of Naples "Parthenope"
Department of Applied Science
Via A. De Gasperi 5, 80131 Naples, Italy
E-mail: alfredo.petrosino@uniparthenope.it

Library of Congress Control Number: 2005938510

CR Subject Classification (1998): I.2.3, I.5, F.4.1, F.1, F.2, G.2, I.2, I.4

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-540-31019-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-31019-8 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: 10983652 06/3142 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 2955
Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 3848: J.-F. Boulicaut, L. De Raedt, H. Mannila (Eds.), Constraint-Based Mining and Inductive Databases. X, 401 pages. 2005.

Vol. 3847: K.P. Jantke, A. Lunzer, N. Spyratos, Y. Tanaka (Eds.), Federation over the Web. X, 215 pages. 2005.

Vol. 3835: G. Sutcliffe, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIV, 744 pages. 2005.

Vol. 3817: M. Faundez-Zanuy, L. Janer, A. Esposito, A. Satue-Villar, J. Roure, V. Espinosa-Duro (Eds.), Nonlinear Analyses and Algorithms for Speech Processing. XII, 380 pages. 2005.

Vol. 3814: M. Maybury, O. Stock, W. Wahlster (Eds.), Intelligent Technologies for Interactive Entertainment. XV, 342 pages. 2005.

Vol. 3809: S. Zhang, R. Jarvis (Eds.), AI 2005: Advances in Artificial Intelligence. XXVII, 1344 pages. 2005.

Vol. 3808: C. Bento, A. Cardoso, G. Dias (Eds.), Progress in Artificial Intelligence. XVIII, 704 pages. 2005.

Vol. 3802: Y. Hao, J. Liu, Y.-P. Wang, Y.-m. Cheung, H. Yin, L. Jiao, J. Ma, Y.-C. Jiao (Eds.), Computational Intelligence and Security, Part II. XLII, 1166 pages. 2005.

Vol. 3801: Y. Hao, J. Liu, Y.-P. Wang, Y.-m. Cheung, H. Yin, L. Jiao, J. Ma, Y.-C. Jiao (Eds.), Computational Intelligence and Security, Part I. XLI, 1122 pages. 2005.

Vol. 3789: A. Gelbukh, Á. de Albornoz, H. Terashima-Marín (Eds.), MICAI 2005: Advances in Artificial Intelligence. XXVI, 1198 pages. 2005.

Vol. 3782: K.-D. Althoff, A. Dengel, R. Bergmann, M. Nick, T. Roth-Berghofer (Eds.), Professional Knowledge Management. XXIII, 739 pages. 2005.

Vol. 3763: H. Hong, D. Wang (Eds.), Automated Deduction in Geometry. X, 213 pages. 2006.

Vol. 3735: A. Hoffmann, H. Motoda, T. Scheffer (Eds.), Discovery Science. XVI, 400 pages. 2005.

Vol. 3734: S. Jain, H.U. Simon, E. Tomita (Eds.), Algorithmic Learning Theory. XII, 490 pages. 2005.

Vol. 3721: A.M. Jorge, L. Torgo, P.B. Brazdil, R. Camacho, J. Gama (Eds.), Knowledge Discovery in Databases: PKDD 2005. XXIII, 719 pages. 2005.

Vol. 3720: J. Gama, R. Camacho, P.B. Brazdil, A.M. Jorge, L. Torgo (Eds.), Machine Learning: ECML 2005. XXIII, 769 pages. 2005.

Vol. 3717: B. Gramlich (Ed.), Frontiers of Combining Systems. X, 321 pages. 2005.

Vol. 3702: B. Beckert (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. XIII, 343 pages. 2005.

Vol. 3698: U. Furbach (Ed.), KI 2005: Advances in Artificial Intelligence. XIII, 409 pages. 2005.

Vol. 3690: M. Pěchouček, P. Petta, L.Z. Varga (Eds.), Multi-Agent Systems and Applications IV. XVII, 667 pages. 2005.

Vol. 3684: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part IV. LXXIX, 933 pages. 2005.

Vol. 3683: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. LXXX, 1397 pages. 2005.

Vol. 3682: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. LXXIX, 1371 pages. 2005.

Vol. 3681: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXXX, 1319 pages. 2005.

Vol. 3673: S. Bandini, S. Manzoni (Eds.), AI*IA 2005: Advances in Artificial Intelligence. XIV, 614 pages. 2005.

Vol. 3662: C. Baral, G. Greco, N. Leone, G. Terracina (Eds.), Logic Programming and Nonmonotonic Reasoning. XIII, 454 pages. 2005.

Vol. 3661: T. Panayiotopoulos, J. Gratch, R.S. Aylett, D. Ballin, P. Olivier, T. Rist (Eds.), Intelligent Virtual Agents. XIII, 506 pages. 2005.

Vol. 3658: V. Matoušek, P. Mautner, T. Pavelka (Eds.), Text, Speech and Dialogue. XV, 460 pages. 2005.

Vol. 3651: R. Dale, K.-F. Wong, J. Su, O.Y. Kwong (Eds.), Natural Language Processing – IJCNLP 2005. XXI, 1031 pages. 2005.

Vol. 3642: D. Ślęzak, J. Yao, J.F. Peters, W. Ziarko, X. Hu (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Part II. XXIII, 738 pages. 2005.

Vol. 3641: D. Ślęzak, G. Wang, M. Szczuka, I. Düntsch, Y. Yao (Eds.), Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Part I. XXIV, 742 pages. 2005.

Vol. 3635: J.R. Winkler, M. Niranjan, N.D. Lawrence (Eds.), Deterministic and Statistical Methods in Machine Learning. VIII, 341 pages. 2005.

Vol. 3632: R. Nieuwenhuis (Ed.), Automated Deduction – CADE-20. XIII, 459 pages. 2005.

Vol. 3630: M.S. Capcarrère, A.A. Freitas, P.J. Bentley, C.G. Johnson, J. Timmis (Eds.), Advances in Artificial Life. XIX, 949 pages. 2005.

Vol. 3626: B. Ganter, G. Stumme, R. Wille (Eds.), Formal Concept Analysis. X, 349 pages. 2005.

Vol. 3625: S. Kramer, B. Pfahringer (Eds.), Inductive Logic Programming. XIII, 427 pages. 2005.

Vol. 3620: H. Muñoz-Ávila, F. Ricci (Eds.), Case-Based Reasoning Research and Development. XV, 654 pages. 2005.

Vol. 3614: L. Wang, Y. Jin (Eds.), Fuzzy Systems and Knowledge Discovery, Part II. XLI, 1314 pages. 2005.

Vol. 3613: L. Wang, Y. Jin (Eds.), Fuzzy Systems and Knowledge Discovery, Part I. XLI, 1334 pages. 2005.

Vol. 3607: J.-D. Zucker, L. Saitta (Eds.), Abstraction, Reformulation and Approximation. XII, 376 pages. 2005.

Vol. 3601: G. Moro, S. Bergamaschi, K. Aberer (Eds.), Agents and Peer-to-Peer Computing. XII, 245 pages. 2005.

Vol. 3596: F. Dau, M.-L. Mugnier, G. Stumme (Eds.), Conceptual Structures: Common Semantics for Sharing Knowledge. XI, 467 pages. 2005.

Vol. 3593: V. Mařík, R. W. Brennan, M. Pěchouček (Eds.), Holonic and Multi-Agent Systems for Manufacturing. XI, 269 pages. 2005.

Vol. 3587: P. Perner, A. Imiya (Eds.), Machine Learning and Data Mining in Pattern Recognition. XVII, 695 pages. 2005.

Vol. 3584: X. Li, S. Wang, Z.Y. Dong (Eds.), Advanced Data Mining and Applications. XIX, 835 pages. 2005.

Vol. 3581: S. Miksch, J. Hunter, E.T. Keravnou (Eds.), Artificial Intelligence in Medicine. XVII, 547 pages. 2005.

Vol. 3577: R. Falcone, S. Barber, J. Sabater-Mir, M.P. Singh (Eds.), Trusting Agents for Trusting Electronic Societies. VIII, 235 pages. 2005.

Vol. 3575: S. Wermter, G. Palm, M. Elshaw (Eds.), Biomimetic Neural Learning for Intelligent Robots. IX, 383 pages. 2005.

Vol. 3571: L. Godo (Ed.), Symbolic and Quantitative Approaches to Reasoning with Uncertainty. XVI, 1028 pages. 2005.

Vol. 3559: P. Auer, R. Meir (Eds.), Learning Theory. XI, 692 pages. 2005.

Vol. 3558: V. Torra, Y. Narukawa, S. Miyamoto (Eds.), Modeling Decisions for Artificial Intelligence. XII, 470 pages. 2005.

Vol. 3554: A.K. Dey, B. Kokinov, D.B. Leake, R. Turner (Eds.), Modeling and Using Context. XIV, 572 pages. 2005.

Vol. 3550: T. Eymann, F. Klügl, W. Lamersdorf, M. Klusch, M.N. Huhns (Eds.), Multiagent System Technologies. XI, 246 pages. 2005.

Vol. 3539: K. Morik, J.-F. Boulicaut, A. Siebes (Eds.), Local Pattern Detection. XI, 233 pages. 2005.

Vol. 3538: L. Ardissono, P. Brna, A. Mitrović (Eds.), User Modeling 2005. XVI, 533 pages. 2005.

Vol. 3533: M. Ali, F. Esposito (Eds.), Innovations in Applied Artificial Intelligence. XX, 858 pages. 2005.

Vol. 3528: P.S. Szczepaniak, J. Kacprzyk, A. Niewiadomski (Eds.), Advances in Web Intelligence. XVII, 513 pages. 2005.

Vol. 3518: T.-B. Ho, D. Cheung, H. Liu (Eds.), Advances in Knowledge Discovery and Data Mining. XXI, 864 pages. 2005.

Vol. 3508: P. Bresciani, P. Giorgini, B. Henderson-Sellers, G. Low, M. Winikoff (Eds.), Agent-Oriented Information Systems II. X, 227 pages. 2005.

Vol. 3505: V. Gorodetsky, J. Liu, V.A. Skormin (Eds.), Autonomous Intelligent Systems: Agents and Data Mining. XIII, 303 pages. 2005.

Vol. 3501: B. Kégl, G. Lapalme (Eds.), Advances in Artificial Intelligence. XV, 458 pages. 2005.

Vol. 3492: P. Blache, E.P. Stabler, J.V. Busquets, R. Moot (Eds.), Logical Aspects of Computational Linguistics. X, 363 pages. 2005.

Vol. 3490: L. Bolc, Z. Michalewicz, T. Nishida (Eds.), Intelligent Media Technology for Communicative Intelligence. X, 259 pages. 2005.

Vol. 3488: M.-S. Hacid, N.V. Murray, Z.W. Raś, S. Tsumoto (Eds.), Foundations of Intelligent Systems. XIII, 700 pages. 2005.

Vol. 3487: J.A. Leite, P. Torroni (Eds.), Computational Logic in Multi-Agent Systems. XII, 281 pages. 2005.

Vol. 3476: J.A. Leite, A. Omicini, P. Torroni, P. Yolum (Eds.), Declarative Agent Languages and Technologies II. XII, 289 pages. 2005.

Vol. 3464: S.A. Brueckner, G.D.M. Serugendo, A. Karageorgos, R. Nagpal (Eds.), Engineering Self-Organising Systems. XIII, 299 pages. 2005.

Vol. 3452: F. Baader, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XI, 562 pages. 2005.

Vol. 3451: M.-P. Gleizes, A. Omicini, F. Zambonelli (Eds.), Engineering Societies in the Agents World V. XIII, 349 pages. 2005.

Vol. 3446: T. Ishida, L. Gasser, H. Nakashima (Eds.), Massively Multi-Agent Systems I. XI, 349 pages. 2005.

Vol. 3445: G. Chollet, A. Esposito, M. Faúndez-Zanuy, M. Marinaro (Eds.), Nonlinear Speech Modeling and Applications. XIII, 433 pages. 2005.

Vol. 3438: H. Christiansen, P.R. Skadhauge, J. Villadsen (Eds.), Constraint Solving and Language Processing. VIII, 205 pages. 2005.

Vol. 3430: S. Tsumoto, T. Yamaguchi, M. Numao, H. Motoda (Eds.), Active Mining. XII, 349 pages. 2005.

Vol. 3419: B.V. Faltings, A. Petcu, F. Fages, F. Rossi (Eds.), Recent Advances in Constraints. X, 217 pages. 2005.

Vol. 3416: M.H. Böhlen, J. Gamper, W. Polasek, M.A. Wimmer (Eds.), E-Government: Towards Electronic Democracy. XIII, 311 pages. 2005.

Vol. 3415: P. Davidsson, B. Logan, K. Takadama (Eds.), Multi-Agent and Multi-Agent-Based Simulation. X, 265 pages. 2005.

Vol. 3413: K. Fischer, M. Florian, T. Malsch (Eds.), Socionics. X, 315 pages. 2005.

Vol. 3403: B. Ganter, R. Godin (Eds.), Formal Concept Analysis. XI, 419 pages. 2005.

Vol. 3398: D.-K. Baik (Ed.), Systems Modeling and Simulation: Theory and Applications. XIV, 733 pages. 2005.

Vol. 3397: T.G. Kim (Ed.), Artificial Intelligence and Simulation. XV, 711 pages. 2005.

Preface

The present volume contains the contributions delivered at the 5th International Workshop on Fuzzy Logic and Applications (WILF 2003), hosted by the Istituto Italiano Studi Filosofici, Palazzo Serra di Cassano, Naples (Italy) and held on October 9-11, 2003.

The volume includes the more recent achievements in the domain of theoretical, experimental and applied fuzzy logic and related techniques. To emphasize the particular connotation of the modern applications of fuzzy logic, special attention has been devoted to the recent trend of integrating and complementing fuzzy logic with rough set theory, neural networks, genetic algorithms and other formal theories and methodologies in order to define flexible and "intelligent" systems, based on the so-called paradigm of soft computing. The capabiblity of these techniques to incorporate imprecision and incomplete information, and to model complex systems, makes them useful tools in many scientific areas.

Among these areas, WILF 2003 dedicated a Special Session on "Soft Computing in Image Processing." Image processing has been a major topic in many areas of research and development, particularly in computer vision and pattern recognition. The majority of the methods were based on probabilistic paradigms, such as the well-known Bayesian paradigm and evidence-based decision-making systems, and just recently soft-computing techniques have gained a relevant role in the leading techniques to tackle image-processing problems. The special session was organized in cooperation with the SCIP group (http://fuzzy.rug.ac.be/SCIP).

The volume consists of peer-reviewed papers, selected out of more than 50 papers submitted to the workshop and given as oral contributions at the workshop. The conference also included three presentations from keynote speakers, Isabelle Bloch from ENST, France, Antonio Di Nola from the University of Salerno, Italy, and Sankar Pal from the Indian Statistical Institute, India.

Thanks are due to Programm Commitee Members and Referees, who took care of the unexpected load of reviewing work. Thanks are also due to the sponsors, with special mention of Antonio Gargano and Gerardo Marotta, director and president of IISF respectively, for supporting the workshop with their financial and organizational help.

Vito Di Gesú, Francesco Masulli and Alfredo Petrosino
Program Chairs
WILF 2003

Organization

WILF 2003 was jointly organized by the Istituto Italiano Studi Filosofici, IISF, the IEEE Neural Networks Society - Italian RIG, the INNS International Neural Network Society, SIG Italy, and SIREN, and by the National Group of Scientific Computing (GNCS), Italy

Executive Committee

Conference Chairs:
Vito Di Gesú (University of Palermo, Italy)
Francesco Masulli (University of Pisa, Italy)
Alfredo Petrosino (University of Naples "Parthenope", Italy)

Program Committee

Jim Bezdek (University of West Florida, USA)
Palma Blonda (CNR-Bari, Italy)
Andrea Bonarini (Politecnico di Milano, Italy)
Piero Bonissone (General Electric, USA)
Ernesto Damiani (University of Milano, Italy)
Antonio Di Nola (University of Salerno, Italy)
Silvio Giove (University of Venezia, Italy)
Marco Gori (University of Siena, Italy)
Ugur Halici (METU, Ankara, Turkey)
Jim Keller (University of Missouri-Columbia, USA)
Etienne Kerre (Ghent University, Belgium)
Ludmilla Kuncheva (University of Wales, UK)
Carlo Morabito (University of Reggio Calabria, Italy)
Gabriella Pasi (CNR-Milano, Italy)
Witold Pedrycz (University of Alberta, Canada)
Roberto Tagliaferri (University of Salerno, Italy)
Settimo Termini (University of Palermo and CNR-Naples, Italy)
Ronald Yager (Iona College, New York, USA)
Hans-Jurgen Zimmermann (RWTH-Aachen, Germany)

Sponsoring Institutions

Istituto Italiano Studi Filosofici (IISF), Naples, Italy
National Group of Scientific Computing (GNCS), Italy
ICAR, National Research Council, Section of Naples, Italy

Table of Contents

Fuzzy Sets and Systems

Fuzzy Control

Neuro-fuzzy Systems

Fuzzy Decision Theory and Application

Soft Computing in Image Processing

Rough-Fuzzy Granular Computing, Case Based Reasoning and Data Mining

Sankar K. Pal

Machine Intelligence Unit,
Indian Statistical Institute, Kolkata, India
sankar@isical.ac.in

Abstract. Data mining and knowledge discovery is described from pattern recognition point of view along with the relevance of soft computing. Key features of the computational theory of perceptions (CTP) and its significance in pattern recognition and knowledge discovery problems are explained. Role of fuzzy-granulation (f-granulation) in machine and human intelligence, and its modeling through rough-fuzzy integration are discussed. Merits of fuzzy granular computation, in terms of performance and computation time, for the task of case generation in large scale case based reasoning systems are illustrated through examples.

Keywords: soft computing, fuzzy granulation, granular computation, rough sets, case based reasoning.

1 Introduction

In recent years, the rapid advances being made in computer technology have ensured that large sections of the world population have been able to gain easy access to computers on account of falling costs worldwide, and their use is now commonplace in all walks of life. Government agencies, scientific, business and commercial organizations are routinely using computers not just for computational purposes but also for storage, in massive databases, of the immense volumes of data that they routinely generate, or require from other sources. Large-scale computer networking has ensured that such data has become accessible to more and more people. In other words, we are in the midst of an information explosion, and there is urgent need for methodologies that will help us bring some semblance of order into the phenomenal volumes of data that can readily be accessed by us with a few clicks of the keys of our computer keyboard. Traditional statistical data summarization and database management techniques are just not adequate for handling data on this scale, and for extracting intelligently, information or, rather, knowledge that may be useful for exploring the domain in question or the phenomena responsible for the data, and providing support to decision-making processes. This quest had thrown up some new phrases, for example, *data mining* [1, 2] and *knowledge discovery in databases (KDD)* which are perhaps self-explanatory, but will be briefly discussed in the following few paragraphs. Their relationship with the discipline of pattern recognition (PR), certain challenging issues, and the role of soft computing will also be mentioned.

V. Di Gesú, F. Masulli, and A. Petrosino (Eds.): WILF 2003, LNAI 2955, pp. 1–10, 2006.

The massive databases that we are talking about are generally characterized by the presence of not just numeric, but also textual, symbolic, pictorial and aural data. They may contain redundancy, errors, imprecision, and so on. KDD is aimed at discovering natural structures within such massive and often heterogeneous data. Therefore PR plays a significant role in KDD process. However, KDD is being visualized as not just being capable of knowledge discovery using generalizations and magnifications of existing and new pattern recognition algorithms, but also the adaptation of these algorithms to enable them to process such data, the storage and accessing of the data, its preprocessing and cleaning, interpretation, visualization and application of the results, and the modeling and support of the overall human-machine interaction. What really makes KDD feasible today and in the future is the rapidly falling cost of computation, and the simultaneous increase in computational power, which together make possible the routine implementation of sophisticated, robust and efficient methodologies hitherto thought to be too computation-intensive to be useful. A block diagram of KDD is given in Figure 1 [3].

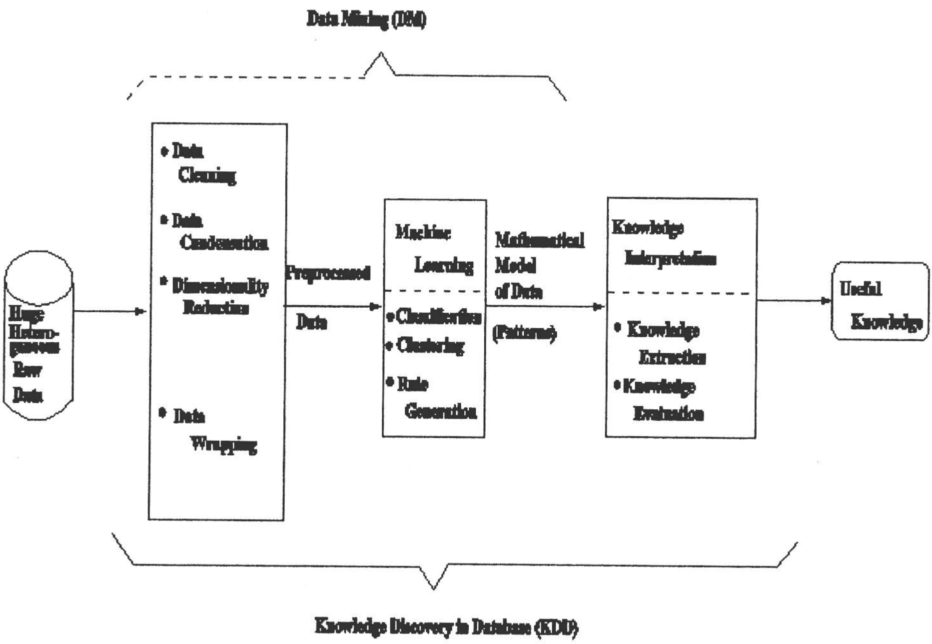

Fig. 1. Block diagram for knowledge discovery in databases [3]

Data mining is that part of knowledge discovery which deals with the process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data, and excludes the knowledge interpretation part of KDD. Therefore, as it stands now, data mining can be viewed as applying PR and machine learning principles in the context of voluminous, possibly heterogeneous data sets. Furthermore, soft computing-based (involving fuzzy sets, neural networks, genetic algorithms and rough sets) PR methodologies and machine learning techniques hold great promise for data mining. The motivation for this is provided by their ability to handle imprecision, vagueness, uncertainty, approximate reasoning and partial truth and lead to tractability, robustness and low-cost solutions [4]. An excellent survey demonstrating the significance of soft computing tools in data mining problem is recently provided by Mitra et al. [5]. Some of the challenges arising out of those posed by massive data and high dimensionality, nonstandard and incomplete data, and over-fitting problems deal mostly with issues like user interaction, use of prior knowledge, assessment of statistical significance, learning from mixed media data, management of changing (dynamic) data and knowledge, integration of different classical and modern soft computing tools, and making knowledge discovery more understandable to humans by using linguistic rules, visualization, etc.

Web mining can be broadly defined as the discovery and analysis of useful information from the web or WWW which is a vast collection of completely uncontrolled heterogeneous documents. Since the web is huge, diverse and dynamic, it raises the issues of scalability, heterogeneity and dynamism, among others. Recently, a detailed review explaining the state of the art and the future directions for web mining research in soft computing framework is provided by Pal et al. [6]. One may note that web mining, although considered to be an application area of data mining on the WWW, demands a separate discipline of research. The reason is that web mining has its own characteristic problems (e.g., page ranking, personalization), because of the typical nature of the data, components involved and tasks to be performed, which can not be usually handled within the conventional framework of data mining and analysis. Moreover, being an interactive medium, human interface is a key component of most web applications. Some of the issues which have come to light, as a result, concern with - (a) need for handling context sensitive and imprecise queries, (b) need for summarization and deduction, and (c) need for personalization and learning. Accordingly, web intelligence became an important and urgent research field that deals with a new direction for scientific research and development by exploring the fundamental roles and practical impacts of machine intelligence and information technology (IT) on the next generation of web-empowered products, systems, services and activities. It plays a key role in today's IT in the era of WWW and agent intelligence.

Bioinformatics which can be viewed as a discipline of using *computational methods to make biological discoveries* [7] has recently been considered as another important candidate for data mining applications. It is an interdisciplinary field mainly involving biology, computer science, mathematics and statistics to analyze biological sequence data, genome content and arrangement, and to predict the function and structure of macromolecules. The ultimate goal is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be derived. There are three major sub-disciplines dealing with the following three tasks in bioinformatics:

a) Development of new algorithms and models to assess different relationships among the members of a large biological data set;
b) Analysis and interpretation of various types of data including nucleotide and amino acid sequences, protein domains, and protein structures; and
c) Development and implementation of tools that enable efficient access and management of different types of information.

First one concerns with the mathematical and computational aspects, while the other two are related to the biological and data base aspects respectively. Data analysis tools used earlier in bioinformatics were mainly based on statistical techniques like regression and estimation. With the need of handling large heterogeneous data sets in biology in a robust and computationally efficient manner, soft computing, which provides machinery for handling uncertainty, learning and adaptation with massive parallelism, and powerful search and imprecise reasoning, has recently gained the attention of researchers for their efficient mining.

While talking about pattern recognition and data mining in the 21st century, it will remain incomplete without the mention of the *Computational Theory of Perceptions (CTP)*, recently explained by Zadeh [8, 9], which has a significant role in the said tasks. In the following section we discuss its basic concepts and features, and relation with soft computing.

2 Computational Theory of Perceptions and F-Granulation

Computational theory of perceptions (CTP) [8, 9] is inspired by the remarkable human capability to perform a wide variety of physical and mental tasks, including recognition tasks, without any measurements and any computations. Typical everyday examples of such tasks are parking a car, driving in city traffic, cooking meal, understanding speech, and recognizing similarities. This capability is due to the crucial ability of human brain to manipulate perceptions of time, distance, force, direction, shape, color, taste, number, intent, likelihood, and truth, among others.

Recognition and perception are closely related. In a fundamental way, a recognition process may be viewed as a sequence of decisions. Decisions are based on information. In most realistic settings, decision-relevant information is a mixture of measurements and perceptions; e.g., the car is six year old but looks almost new. An essential difference between measurement and perception is that in general, measurements are crisp, while perceptions are fuzzy. In existing theories, perceptions are converted into measurements, but such conversions in many cases, are infeasible, unrealistic or counterproductive. An alternative, suggested by the CTP, is to convert perceptions into propositions expressed in a natural language, e.g., it is a warm day, he is very honest, it is very unlikely that there will be a significant increase in the price of oil in the near future.

Perceptions are intrinsically imprecise. More specifically, perceptions are f-granular, that is, both fuzzy and granular, with a granule being a clump of elements of a class that are drawn together by indistinguishability, similarity, proximity or functionality. For example, a perception of height can be described as very tall, tall, middle, short, with very tall, tall, and so on constituting the granules of the variable 'height'. F-granularity of perceptions reflects the finite ability of sensory organs and, ultimately, the brain, to resolve detail and store information. In effect, f-granulation is a human way of achieving data compression. It may be mentioned here that although information granulation in which the granules are crisp, i.e., c-granular, plays key roles in both human and machine intelligence, it fails to reflect the fact that, in much, perhaps most, of human reasoning and concept formation the granules are fuzzy (f-granular) rather than crisp. In this respect, generality increases as the information ranges from singular (age: 22 yrs), c-granular (age: 20-30 yrs) to f-granular (age: "young"). It means CTP has, in principle, higher degree of generality than qualitative reasoning and qualitative process theory in AI [10, 11]. The types of problems that fall under the scope of CTP typically include: perception based function modeling, perception based system modeling, perception based time series analysis, solution of perception based equations, and computation with perception based probabilities where perceptions are described as a collection of different linguistic *if-then* rules.

F-granularity of perceptions puts them well beyond the meaning representation capabilities of predicate logic and other available meaning representation methods. In CTP, meaning representation is based on the use of so called constraint-centered semantics, and reasoning with perceptions is carried out by goal-directed propagation of generalized constraints. In this way, the CTP adds to existing theories the capability to operate on and reason with perception-based information.

This capability is already provided, to an extent, by fuzzy logic and, in particular, by the concept of a linguistic variable and the calculus of fuzzy if-then rules. The CTP extends this capability much further and in new directions. In application to pattern recognition and data mining, the CTP opens the door to a much wider and more systematic use of natural languages in the description of patterns, classes, perceptions and methods of recognition, organization, and knowledge discovery. Upgrading a search engine to a question- answering system is another prospective candidate in web mining for CTP application. However, one may note that dealing with perception-based information is more complex and more effort-intensive than dealing with measurement-based information, and this complexity is the price that has to be paid to achieve superiority.

3 Granular Computation and Rough-Fuzzy Approach

Rough set theory [12] provides an effective means for analysis of data by synthesizing or constructing approximations (upper and lower) of set concepts from the acquired data. The key notions here are those of "information granule" and "reducts". Information granule formalizes the concept of finite precision representation of objects in real life situation, and reducts represent the core of an information system (both in terms of objects and features) in a granular universe. *Granular computing* refers to that where computation and operations are performed on information granules (clump of similar objects or points). There-