Cognitive Informatics and Computational Intelligence: Theory and Applications

Preface

This special issue of Fundamenta Informaticae focuses on the foundations and applications of Cognitive Informatics and Computational Intelligence (briefly, CI2). CI2 focuses on studies of human information processing as well as the byproducts of perception and cognition. Cognitive Informatics (CI) is a multidisciplinary study of cognition, computing and information sciences which investigates the information processing mechanisms and processes of the brain and natural intelligence, as well as their engineering applications in cognitive computing.

Specifically, CI2 provides a coherent set of fundamental theories and contemporary mathematics that form the foundation for most science and engineering disciplines such as applied mathematics (e.g., perceptual forms of fuzzy sets, near sets and rough sets), computer science, cognitive science, computer engineering (e.g., computer vision), cybernetics (e.g., machine behavior), neuropsychology and pure mathematics (e.g., proximity spaces, topological spaces via near and far).

This special issue presents some of the latest advances in cognitive informatics and cognitive computing. A total of 11 papers were accepted for publication. Each accepted paper has undergone a thorough review (at least two reviewers for each paper) and a second round of review and revision cycle.

The paper by M.H-Herrero, P. Rabanal, I. Rodríguez, and F. Rubio on *Comparing Problem Solving Strategies for NP-hard Optimization Problems*, present analysis of performance of humans when solving NP-complete problems. These analyses are supported by experiments which include the human capability to compute good suboptimal solutions to these problems, and the authors try to identify the kind of problem instances which make humans compute the best and worst solutions (including the dependance of their performance on the size of problem instances). Finally, their performance with computational heuristics typically used to approximately solve these problems are compared, and participants in these experiments are also interviewed in order to infer the most typical strategies used by them, as well as how these strategies depend on the form and size of problem instances.

The paper by G. Virginia and H.S. Nguyen on *Lexicon-based Document Representation*, is based on tolerance rough sets model(TRSM) to model document-term relations in text mining. Specifically, this representation maps the terms occurring in TRSM-representation to terms in the lexicon, hence the final representation of a document is a weight vector consisting only of terms that occurred in the lexicon (lexicon-representation). The lexicon-representation can be viewed as a compact form of TRSM-

representation in a lower dimensional space and eliminates all informal terms previously occurring in TRSM-vector. The effectiveness of lexicon-representation is compared with TRSM-representation using recall and precision measures.

The paper by E. R. Hruschka Jr., M. C. Duarte and M. C. Nicoletti on *Coupling as Strategy for Reducing Concept-Drift in Never-ending Learning Environments* formalize relevant coupling strategies for dealing with the concept-drift problem in a Never-ending Learning (NEL) environment implemented as the system RTWP (Read The Web in Portuguese). Bootstrap learning has been one of the most successful among the semi-supervised methods proposed to date and, as such, the natural candidate for implementing NEL systems. Initial results have shown they are promising strategies for minimizing the problem taking into account a few system settings. Considering the principles of never-ending learning environments, precision is a key factor to prevent NEL systems to drift from what it was initially intended.

In the paper, *Unsupervised Tracking, Roughness and Quantitative Indices* by Sankar K. Pal and Debarati Chakraborty, a novel methodology for tracking a single moving object in a video sequence applying the concept of rough set theory is presented. The novelty of this technique is that it does not consider any prior information about the video sequence unlike many existing techniques. The first target model is constructed using the median filtering based foreground detection technique and after that the target is reconstructed in every frame according to the rough set based feature reduction concept incorporating a measure of indiscernibility instead of indiscernibility matrix. Three quantitative indices based on rough sets, feature similarity and Bhattacharya distance are proposed to evaluate the performance of tracking and detect the mis-tracked frames in the process of tracking to make those corrected. Extensive experimental results are given to demonstrate the effectiveness of the method. Comparative performance is demonstrated both visually and quantitatively.

The paper, *"Knowledge Representation for Human-Computer Interaction in a System Supporting Conceptual Design*, Ewa Grabska, Grażyna Ślusarczyk and Szymon Gajek deals with the conceptual stage of the design process related to civil engineering. Different types of design knowledge representation that are essential to visual aspects of a human computer dialogue are considered. The design knowledge includes: drawings expressing forms, layouts and functionality of designed artifacts, internal graph-based data structures obtained automatically on the basis of these drawings and logic formulas extracted from the graph data structures. The reasoning mechanism based on the first-order logic enables the system to assess the compatibility of designs with specified requirements and constraints. The feedback given by the system along with visualizations of initial design ideas enforce the designer's constructive visual perception and creativity. The presented approach is illustrated using the example of designing an indoor swimming pool.

The paper by J. H. Saito, J.-F. Mari, E. Pedrino, J. B. Destro Filho and M. C. Nicoletti on *Simulated Activation Patterns of Biological Neurons Cultured onto a Multi-Electrode Array Based on a Modified Izhikevichi's Model* discusses Sim-MEA, a system for modeling and simulating neuron's communications in a MEA-based in vitro culture. Sim-MEA implements a modified Izhikevich model that takes into account both distances between neurons and distances between microelectrodes and neurons. The system also provides ways of simulating microelectrodes and their recorded signals as well as recovering experimental MEA culture data from their images. The soundness of the Sim-MEA simulation procedure was empirically evaluated using data from experimental in vitro cultured hippocampal neurons of Wistar rat embryos. Results from simulations are compared with those of the in-vitro experiment are presented and discussed. The paper also describes a few experiments with varying parameter values to illustrate the approach.

In the paper, *Automatic Learning of Temporal Relations under the Closed World Assumption* M. C. Nicoletti, F. O. S. S. Lisboa and E. R. Hruschka Jr explore the automatic learning of Allen's temporal relations by the inductive logic programming system FOIL, taking into account two possible representations for a time interval: (i) as a primitive concept and (ii) as a concept defined by the primitive concept of time point. The goals of the experiments described in the paper are (1) to explore the viability of both representations for use in automatic learning; (2) to compare the facility and interpretability of the results; (3) to evaluate the impact of the given examples for inducing a proper representation of the relations and (4) to experiment with both representations under the assumption of a closed world (CWA), which would ease continuous learning using FOIL. Experimental results are presented and discussed as evidence that the CWA can be a convenient strategy when learning Allen's temporal relations.

In this paper titled *Robust Rough-Fuzzy C-Means Algorithm: Design and Applications in Coding and Non-coding RNA Expression Data Clustering*, Pradipta Maji and Sushmita Paul present a new rough-fuzzy clustering algorithm, termed as robust rough-fuzzy c-means. Each cluster in the proposed clustering algorithm is represented by a set of three parameters, namely, cluster prototype, a possibilistic fuzzy lower approximation, and a probabilistic fuzzy boundary. The possibilistic lower approximation helps in discovering clusters of various shapes. The cluster prototype depends on the weighting average of the possibilistic lower approximation and probabilistic boundary. The proposed algorithm is robust in the sense that it can find overlapping and vaguely defined clusters with arbitrary shapes in noisy environment. An efficient method is presented, based on Pearson's correlation coefficient, to select initial prototypes of different clusters. A method is also introduced based on cluster validity index to identify optimum values of different parameters of the initialization method and the proposed clustering algorithm. The effectiveness of the proposed algorithm, along with a comparison with other clustering algorithms, is demonstrated on synthetic as well as coding and non-coding RNA expression data sets using some cluster validity indices.

In the paper, *Sufficiently Near Neighbourhoods of Points in Flow Graphs – A Near Set Approach*, James F. Peters and Doungrat Chitcharoen introduce sufficiently near visual neighbourhoods of points and neighbourhoods of sets in digital image flow graphs (NDIFGs). An NDIFG is an extension of a Pawlak flow graph. The study of sufficiently near neighbourhoods in NDIFGs stems from recent work on near sets and topological spaces via near and far, especially in terms of visual neighbourhoods of points that are sufficiently near each other. From a topological perspective, non-spatially near sets represent an extension of proximity space theory and the original insight concerning spatially near sets by F. Riesz at the International Congress of Mathematicians (ICM) in 1908. In the context of Herrlich nearness, sufficiently near neighbourhoods of sets in NDIFGs provide a new perspective on topological structures in NDIFGs. The practical implications of this work are significant. With the advent of a study of the nearness of open as well as closed neighbourhoods, it is now possible to do information mining on a more global level and achieve new insights concerning the visual information embodied in the images that provide input to an NDIFG.

In the paper, *An Insight Into The Z-number Approach To CWW*, Sankar K. Pal, Romi Banerjee, Soumitra Dutta and Samar Sen Sarma present a comprehensive investigation of the Z-number approach to CWW by : a) providing an outline of the generic architecture, algorithm and challenges underlying CWW in general; b) performing a detailed study of the Z-number methodology, including an algorithm for CWW using Z-numbers, a Z-number based operator for the evaluation of the level of requirement satisfaction, and simulation experiments of CWW utilizing Z-numbers; and c) analysing the strengths and the challenges of the Z-numbers, and suggest possible solution strategies. The goal is to inspire

research on the need for inclusion of human-behavioural aspects into CWW, as well as the integration of CWW and NLP.

In 1999, Molodtsov introduced the theory of soft sets, which can be seen as a new mathematical approach to vagueness. In 2002, near set theory was initiated by J. F. Peters as a generalization of Z. Pawlak's rough set theory. In the near set approach, every perceptual granule is a set of objects that have their origin in the physical world. Objects that have, in some degree, affinities are considered perceptually near each other, i.e., objects with similar descriptions. In the paper, "Soft Nearness Approximation Spaces", M. Ali Ozturk and E. Inan aim to combine the soft sets approach with near set theory, which gives rise to the new concepts of soft nearness approximation spaces (SNAS), soft lower and upper approximations. Some examples and properties of these soft nearness approximations are given.

Special issue editors

Zhihua Cui
Complex System and Computational Intelligence Laboratory
School of Computer Science & Technology
Taiyuan University of Science and Technology, China, 030024
cuizhihua@gmail.com

Sheela Ramanna
Department of Applied Computer Science
University of Winnipeg, Winnipeg, Manitoba, R3B 2E9, Canada
s.ramanna@uwinnipeg.ca

James F. Peters
Computational Intelligence Laboratory
Department of Electrical & Computer Engineering
University of Manitoba, Winnipeg, Manitoba R3T 5V6, Canada
james.peters3@ad.umanitoba.ca

Sankar K. Pal
Machine Intelligence Unit
Indian Statistical Institute
Kolkata 700 108, India
sankar@isical.ac.in