
Preface

Medical informatics is an emerging interdisciplinary science that deals with clinical health-related information, its structure, acquisition and use. Medical Informatics (MI) is grounded in the principles of computer science, artificial intelligence, as well as the clinical and basic sciences. MI includes scientific endeavors ranging from building theoretical models and evaluation of applied systems.

Computational intelligence is a well-established paradigm, where new theories with a sound biological understanding have been evolving. Defining computational intelligence is not an easy task. In a nutshell, which becomes quite apparent in light of the current research pursuits, the area is heterogeneous with a combination of such technologies as neural networks, fuzzy systems, rough set, evolutionary computation, swarm intelligence, probabilistic reasoning, multi-agent systems etc. The recent trend is to integrate different components to take advantage of complementary features and to develop a synergistic system.

This book deals with the application of computational intelligence in medical informatics. Addressing the various issues of medical informatics using different computational intelligence approaches is the novelty of this edited volume. This volume comprises of 15 chapters' including an introductory chapter giving the fundamental definitions and some important research challenges. Chapters were selected on the basis of fundamental ideas/concepts rather than the thoroughness of techniques deployed. The fifteen chapters are organized as follows.

In the introductory Chapter, *Kelemen et al.* provide a review of recently developed theories and applications in computational intelligence for gene-gene and gene-environment interactions in complex diseases in genetic association study.

Chapter 2 by *Burns et al.* is designed to act as an introduction to the field of biomedical text-mining for computer scientists who are unfamiliar with the way that biomedical research uses the literature. Authors describe how creating knowledge bases from the primary biomedical literature is formally

equivalent to the process of performing a literature review or a ‘research synthesis’. The main body of the chapter is concerned with the use of text mining approaches to populate knowledge representations for different types of experiment. Authors provide a detailed example from neuroscience and describe a detailed description of the methodology used to perform the text mining based on the conditional random fields model.

Tamalika Chaira and *Tridib Chaira* in Chapter 3 propose a new image segmentation technique using intuitionistic fuzzy set and is applied for brain images and blood cell for edge detection. The proposed method works well even on poor quality images.

In Chapter 4, *Kreinovich* and *Shpak* illustrate that in general, detecting aggregability is NP-hard even for linear systems, and thus (unless $P=NP$), there is only hope to find efficient detection algorithms for specific classes of systems. Authors illustrate that in the linear case, once the blocks are known, it is possible to efficiently find appropriate linear combinations.

Siebel et al. in the fifth Chapter propose the automatic design of neural networks as a controller in a visuo-motor control scenario. Evolutionary Acquisition of Neural Topologies (EANT) uses evolutionary search methods on two levels: In an outer optimization loop called structural exploration new networks are developed by gradually adding new structures to an initially minimal network. In an inner optimization loop called structural exploitation the parameters of current networks are optimized. EANT was used with a complete simulation of a visuo-motor control scenario to learn neural networks by reinforcement learning.

In Chapter 6, *Lee et al.* propose Block Principal Components Analysis (PCA) and a variable selection method based on principal component loadings for dimension reduction. The main focus of this is how to deal with large number of variables (gene expressions) in microarray data sets. Authors also investigate the effect ill-conditioning has on the Mahalanobis distance between clusters using the well-known Hilbert matrix.

De Roberto Jr et al. in the seventh Chapter describes the development of a new tool for genome interpretation. The software recognizes coding regions with a user-friendly interface. The system is based on a gene model and combines the weight-position matrix technique with the flexibility of artificial neural networks in classification problems.

In Chapter 8, *Bogan-Marta et al.* discuss about the diversity of language engineering techniques and those involving information theoretic principles in analyzing protein sequences from similarity perspective. Authors also present a survey of the different approaches identified with a focus on two methods, which they they experimented.

Sehgal et al. in Chapter 9, investigate the impact of missing values on post genomic knowledge discovery methods like, gene selection and Gene Regulatory Network (GRN) reconstruction. A framework for robust subsequent biological knowledge inference is proposed, which has shown significant improvements in the outcomes of gene selection and GRN reconstruction methods.

In Chapter 10, *Sehgal et al.* provide a comprehensive comparative study on GRN reconstruction algorithms. The methods discussed are diverse and vary from simple similarity based methods to state of the art hybrid and probabilistic methods. The Chapter also emphasizes the need of strategies, which should be able to model the stochastic behavior of gene regulation in the presence of limited number of samples, noisy data, multi-collinearity for high number of genes.

Kasabov et al. in Chapter 11 present some preliminary results on the Brain-Genes Ontology project that is concerned with the collection, presentation and use of knowledge in the form of ontology equipped with the knowledge discovery means of computational intelligence. Brain-Genes Ontology system includes various concepts, facts, data, graphs, visualizations, animations, and other information forms, related to brain functions, brain diseases, their genetic basis and the relationship between all of them, and various software simulators.

In Chapter 12, *Dohnal et al.* present the metric space approach and its applications in the field of Bioinformatics. Authors describe some of the most popular centralized disk-based metric indexes with a focus on parallel and distributed access methods, which can deal with data collections that for practical purposes can be arbitrary large. An experimental evaluation of the presented distributed approaches on real-life data sets is also presented.

Kroc in Chapter 13 illustrates that the development of an adequate mechanical model of living tissues provides the morphological model with sufficient flexibility necessary to achieve expected morphological development. Author focuses on the development of mesenchymal and epithelial tissues, which creates the basic mechanism of tooth development.

In Chapter 14, *Bobrik et al.* describe two simplified models of Darwinian evolution at the molecular level by applying the methods of artificial chemistry. A metaphor of a chemical reactor (chemostat) is considered and then a simplified formal system *Typogenetics*, is discussed.

In the last Chapter *Bobrik et al.* present the third simplified model of Darwinian evolution at the molecular level, following the two presented in Chapter 14. An artificial-life application is designed as a modification of the metaphor of chemostat, where the secondary structure of binary strings specifies instructions for replication of binary strings.

We are very much grateful to the authors of this volume and to the reviewers for their tremendous service by critically reviewing the chapters. The editors would like to thank Dr. Thomas Ditzinger (Springer Engineering Inhouse Editor) and Professor Janusz Kacprzyk (Editor-in-Chief, Springer Studies in Computational Intelligence Series) and Ms. Heather King (Springer Verlag, Heidelberg) for the editorial assistance and excellent cooperative collaboration to produce this important scientific work. We hope that the reader will share our excitement to present this volume on '*Computational Intelligence in Medical Informatics*' and will find it useful.

August 2007

Arpad Kelemen
Ajith Abraham
Yulan Liang