

# Preface

This book is a collection of papers written for the International Workshop *Clusters, orders, trees: Methods and applications* on the occasion of the 70th Anniversary of Boris Mirkin, Professor at National Research University Higher School of Economics (NRU HSE), Moscow Russia, and Birkbeck College University of London, UK, which was held in NRU HSE Moscow, 12 and 13 December 2012, with the following schedule.

## *12 December 2012*

11.00–11.20 Opening

11.20–12.00 Fuad Aleskerov (NRU HSE, Moscow, Russia). Interval orders, semiorders, and their numerical representation.

12.00–12.30 Katya Chernyak (NRU HSE, Moscow, Russia). Scoring extent of similarity between a string and text using suffix tree: method and applications.

12.30–13.00 Yulia Veselova (NRU HSE, Moscow, Russia). The manipulability index in the IANC model.

13.00–15.00 Break

15.00–15.40 Fred Roberts (Rutgers University, NJ, USA). Meaningless statements in landscape ecology and sustainable environments.

15.40–16.20 Vladimir Makarenkov (Université de Québec, Montréal, Canada). Building trees encompassing horizontal transfer events: applications in evolution and linguistics.

16.20–16.50 Break

16.50–17.30 Trevor Fenner (University of London, UK). Lifting algorithm in a tree and its applications.

17.30–18.00 Andrey Shestakov (NRU HSE, Moscow, Russia). Least squares consensus clustering.

## *13 December*

11.00–11.40 Valery Kalyagin (NRU HSE, Nizhniy Novgorod, Russia). Mining market data: Russian stock market.

11.40–12.20 Sergei Archangelski, Ilya Muchnik (Rutgers University, NJ, USA). Clustering in registration of 3D point clouds.

12.20–13.00 Ito Wasito (University of Indonesia). Least-squares imputation of missing data entries.

13.00–14.30 Break

14.30–15.10 Panos Pardalos (University of Florida, USA). High dimensional data classification.

15.10–15.50 Fred McMorris (IIT, Chicago, USA). The majority decision function on median semilattices.

15.50–16.10 Break

16.10–16.50 Susana Nascimento (Universidade Nova de Lisboa, Portugal). Data recovery fuzzy clustering: proportional membership and additive spectral methods.

16.50–17.30 Boris Goldengorin (NRU HSE, Moscow, Russia). Mixed tools for market analysis and their applications

This workshop has reflected some past and present Boris Mirkin's research activities. Boris Mirkin has contributed to the development of all three concepts with outstanding contributions.

In clustering, among others are

1. deriving and using distance between partitions (1969, in Russian) predating the popular Rand index (1971) which is the complement of Mirkin's distance to unity,
2. proposing the so-called qualitative fuzzy analysis model and methods (1976, in Russian) which includes the popular additive clustering model by Shepard and Arabie (1979) and an effective one-by-one approach for identifying it, and
3. principal cluster analysis model and method (1987, in Russian) later converted by him into the Separate-and-Conquer strategy (1998) and then to Anomalous Pattern clustering and Intelligent K-Means methods (2005).

A joint paper by R. Amorim and B. Mirkin in this volume relates to a further extension of the latter approach to Minkowski distances and weighted variables. A joint paper by Mirkin and Shestakov shows how effective can be an approach to consensus clustering outlined by B. Mirkin and I. Muchnik back in 1981 (in Russian), in comparison to recent approaches.

In ordering, among others are:

1. a characterization of the interval orders as those irreflexive relations whose images are linearly ordered over settheoretic inclusion (1969, in Russian), and
2. extension of the Arrow consensus between rankings approach to equivalence relations (1971) and then any relations (1973, in Russian, 1979 in English), and using Arrow axioms for characterization of what he, and then the others, calls federation consensus rules (1979, in Russian).

In trees, among others are:

1. a biologically meaningful mapping of gene phylogenetic trees onto a species evolutionary tree (Mirkin–Muchnik–Smith model 1995) and
2. split vector bases corresponding to binary hierarchies (1996).

A paper by K. Chernyak and B. Mirkin describes further developments in text analysis based on the usage of two tree-related concepts, suffix tree, and taxonomy. It should be pointed out that the concepts of cluster, order, and tree are in the core of all efforts in data analysis, and these are to remain in the core for the foreseeable future, because they are in the core of any structuring attempts.

Of course, no real-world application can be developed without measuring features and relations, and B. Mirkin has contributed to this, as well.

Examples include:

1. matrix-based correlation and association measures in the space of mixed scale variables (1982, in Russian) and
2. least-squares methodology for imputation of missing values. Several papers in this volume can be attributed to this direction.

Accordingly, we divided all the contributions in three sections:

- (a) classification and cluster,
- (b) order and tree, and
- (c) measurement.

In addition to Boris Mirkin's startling Ph.D. results in mathematical logic and algebra, Mirkin's groundbreaking contributions in various fields of decision-making theory and practice have marked the fourth quarter of the twentieth century and beyond. Boris has done pioneering work in group choice, mathematical psychology, clustering, data mining and knowledge discovery which are activities oriented towards finding nontrivial or hidden patterns in data collected in databases. Boris Mirkin has published several books, such as: *The Group Choice Problem* (in Russian, 1974), *Analysis of Quality Indicators* (in Russian, 1976), *Graphs and Genes* (in Russian, co-authored with S.N. Rodin, 1977), *Group Choice* (Wiley-Interscience, 1979), *Analysis of Quality Indicators and Structures* (in Russian, 1976), *Graphs and Genes* (Springer, co-authored with S.N. Rodin, 1984), *Clusters in Social- Economics Research* (in Russian, 1985), *Mathematical Classification and Clustering* (Kluwer, 1996), *Clustering for Data Mining: A Data Recovery Approach* (Chapman and Hall/CRC, 2005; Second Edition, 2012), *Core Concepts in Data Analysis: Summarization, Correlation and Visualization* (Undergraduate Topics in Computer Science) (Springer, 2011).

Our special thanks to all reviewers who made a crucial contribution to the scheduled production of this volume. Here we would like to list all of them: Fyad Aleskerov, Rozenn Dahyot, Anuška Ferligoj, Boris Goldengorin, Dmitry Ignatov, Friedrich Leisch, Vladimir Makarenkov, Boris Mirkin, Sergei Obiedkov, Panos M. Pardalos, Niel J le Roux, and Yulia Veselova.

This volume contains the collection of papers reflecting many developments in theory and applications rooted by Boris' fundamental contribution to the state of the art in group choice, mathematical psychology, clustering, data mining, and knowledge discovery. Researches, students, and engineers will benefit from new knowledge discovery techniques.

Moscow, Russia  
Groningen, The Netherlands  
Gainesville, FL, USA

Fuad Aleskerov  
Boris Goldengorin  
Panos M. Pardalos



<http://www.springer.com/978-1-4939-0741-0>

Clusters, Orders, and Trees: Methods and Applications  
In Honor of Boris Mirkin's 70th Birthday  
Aleskerov, F.T.; Goldengorin, B.; Pardalos, P. (Eds.)  
2014, XI, 404 p. 80 illus., 28 illus. in color., Hardcover  
ISBN: 978-1-4939-0741-0