ifcs

International Federation of Classification Societies Newsletter

Number 19 April 2000

President: Jean-Paul Rasson Editor: Paul De Boeck

The International Federation of Classification Societies, founded in 1985, is composed of:

Associaçáo Portuguesa de Classificaçào e Análise de Dados, British Classification Society, Classification Society of North America, Gesellschaft für Klassifikation, Japanese Classification Society, Korean Classification Society, Société Francophone de Classification, Società Italiana di Statistica, Vereniging voor Ordinatie en Classificatie, and Section of Classification and Data Analysis of the Polish Statistical Society.

The IFCS is a non-profit, non-political scientific organization, the aims of which are to further classification research. Among other activities, the IFCS organises a biennal conference, and supports the Journal of Classification.

A new society has joined the IFCS: the Irish Pattern Recognition and Classification Society, see p. 4

Contents:

From the President

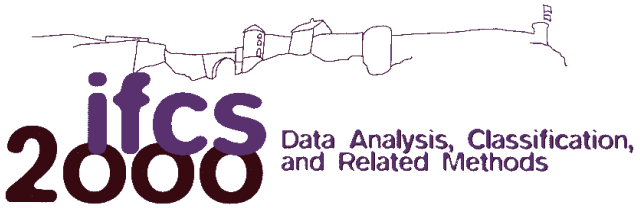
It is a pleasure to thank and to congratulate our prestigious former President, Prof. Chikio Hagashi. He has been very efficient in conducting and leading the IFCS and I think that his main quality was his generosity. To mention only one example, he built and provided the funds for the IFCS Travel Awards Program (TAP) that, as I have recently seen, is very successful amongst the young researchers. They consider it a great honor to be awarded a TAP grant.

As suggested by some members of the executive committee, I would like, in a few words, to introduce myself: I am full professor and director of the department of mathematics of the "Facultés Universitaires Notre-Dame de la Paix" at Namur (Belgium). I prepared my Ph.D. thesis in Paris (Université Paul Descartes) and at the University of Cambridge (U.K.) and I received the Ph.D. degree in 1978 (the topics of the thesis were Probability and Statistics). My research is mainly concerned with the geometrical aspects of statistics, pattern recognition and classification with, as favorite fields of application: image analysis and credit-scoring. I have been president of the SFC (Société Francophone de Classification) from 1996 to 1998.

For the near future, the IFCS will face a problem of growth: the Irish pattern recognition and classification society has joined the IFCS last year and, as far as I know, some other new societies will do the same very soon. This is a very good sign of the health of the IFCS, but this is also a bit challenging: we will probably have to adapt the by-laws, at least the way of working of the Council.

I really hope that I will have the pleasure to welcome most of you at our IFCS-2000 meeting in July!

Jean-Paul Rasson



7th Conference of the International Federation of Classification Societies, July 11-14, 2000, Namur, Belgium

The IFCS Biennial Conference

Six IFCS conferences have been held: at Aachen, Germany (1987), Charlottesville, USA (1989), Edinburgh, Scotland (1991), Paris, France (1993), Kobe, Japan (1996), and Rome, Italy (1998).

The next IFCS conference will be held in Namur, Belgium from 11 - 14 July 2000. Further information can be obtained from:

IFCS-2000 Data Analysis, Classification and Related Methods
A. Hardy, Secretary Local Organizing Committee, Department of Mathematics
Namur University Notre-Dame de la Paix
Rempart de la Vierge, 8, B-5000
Namur, Belgium
phone: +32 81 72 49 05
fax: +32 81 72 49 14
email: Andre.Hardy@fundp.ac.be
WWW: http://www.fundp.ac.be/~ifcs2000/

Keynote Speakers

- Andreas Buja (AT&T Labs)
  Data Mining Criteria for Tree-Based Regression and Classification
- Bradley Efron (Stanford University)
  The cost of model selection
- David Hand (Imperial College, London)
  Breaking the mould: some new problems in supervised classification
- John A. Hartigan (Yale University)
  Classifier Probabilities

There will be no further information sent automatically to the registered participants, everything relevant appears on the website (see above).

If you want further information, don't hesitate to contact Andre Hardy (see above).

Two related activities scheduled in July 2000 in Belgium

1. **International Symposium on Consensus Methods and Applications**
   Namur, Belgium July 10, 2000
2. **Ordinal and Symbolic Data Analysis**
   OSDA 2000
   to be held in Brussels, July 5 to 8, 2000

---

1. International Symposium on Consensus Methods and Applications

Sponsored by:
Natural Sciences and Engineering Research Council of Canada
Hosted by:
University of Namur, Belgium
Organizing committee:
Guy Cucumel (Canada), Allan D. Gordon (Scotland), François-Joseph Lapointe (Canada), Maurizio Vichi (Italy)

Theme and goals

This one-day symposium will be held prior to the IFCS-2000 meeting in Namur to promote communication among researchers from different disciplines and countries and to enhance dissemination of current knowledge about consensus methods and applications. This meeting will consist of a mix of invited lectures, contributed papers, and panel discussions. Major topics concern (but are not limited to) consensus methods, algorithmic aspects, supertree methods, validation of consensus, multiple consensus representations, and consensus graphs with applications to biology, sociology, economy, marketing, psychometry, linguistics and engineering, among others.

Call for papers

Abstracts of contributed papers, including the names, affiliations, and email addresses of the authors, should be sent via email or mailed (with diskette) to:

François-Joseph Lapointe
Département de sciences biologiques,
Université de Montréal,
C.P. 6128, Succursale Centre-ville,
Montréal (Québec), Canada, H3C 3J7.
E-mail: lapoinf@ere.umontreal.ca
Fax : 1 514 343 2293
Phone : 1 514 343 7999

The DEADLINE for submitting abstracts is March 31, 2000.

2. Ordinal and Symbolic Data Analysis

Invited speakers

- Michel CHEIN and Marie-Laure MUGNIER
  (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier, CNRS and Université Montpellier 2)
  Graph Homomorphisms: a Tool for Managing Graph Data Bases;
- Andreas DRESS
  (Fakultät für Mathematik, Universität Bielefeld):
  An Excluded-Minor Characterization of Universally Representable Nested Families of Preference Relations
- Jean-Claude FALMAGNE
  (Department of Cognitive Sciences, University of California, Irvine):
  Recent Developments in Media Theory, with Applications;
- Bernhard GANTER
  (Institut für Algebra, Fakultät Mathematik und Naturwissenschaften, Technische Universität Dresden):
  Contextual Logic for Conceptual Knowledge
- Pierre HANSEN
  (GERAD and Ecole des Hautes Etudes Commerciales de Montréal):
  Finding Relations in Graph Theory by Computer (joint work with Gilles Caporossi)
- Itzhak GILBOA
  (The Eitan Berglas School of Economics and The Recanati School of Business, Tel-Aviv University):
  Inductive Inference: An Axiomatic Approach
- R.Duncan LUCE
  (Institute for Mathematical Behavioral Sciences, University of California, Irvine):
  Title to be announced
- William T. TROTTER
  (Department of Mathematics, Arizona State University):
  Partially Ordered Sets: Combinatorial Properties and their Applications
- Philippe VINCKE
  (Service de Mathématiques de la Gestion, Université Libre de Bruxelles):
  Handling Ordinal Data in Multicriteria Decision Analysis: a Survey
- Uta WILLE
  (Data Management, Jelmoli AG):
  Exploring and Communicating Customer Data with TOSCANA

A special session will be organized by Edwin DIDAY in the area of symbolic data analysis.

The conference is scheduled for:
Wednesday July, 5 to Saturday July, 8 of the year 2000

This is just a few days before the conference in Namur, Belgium, of the International Federation of Classification Societies, which will last from Tuesday 11 to Friday 14; see:

http://www.fundp.ac.be/~ifcs2000/ifcs_01.html

Notice also that a one-day meeting on consensus theory will be organized in Namur on Monday July 10 by Francois-Joseph Lapointe
lapoinf@ERE.UMontreal.CA

Moreover, the European Mathematical Society has its 3rd European Congress of Mathematics from July 10 to July 14 in Barcelona; see:

http://www.iec.es/3ecm/

Location
The campus of the Université Libre de Bruxelles (ULB), not far away from the center of Brussels, in the heart of Europe.

Local organizers
Jean-Paul Doignon (ULB), Samuel Fiorini (ULB), Marc Pirlot (Faculté Polytechnique de Mons).

Website
General information on the conference will be posted on a website maintained by Samuel Fiorini at the following URL:

http://www.ulb.ac.be/sciences/ulbmath/osda2000

History and theme
This conference follows five previous conferences which were as follows:

- Conference on Ordinal Data Analysis, TH Darmstadt, March 1992;
- Conference on Ordinal Data Analysis, University of Massachusetts (Amherst), October 1993;
- Conference on Ordinal and Symbolic Data Analysis, Ecole Nationale Supérieure des Télécommunications (Paris), June 1995;
- Ordinal and Symbolic Data Analysis Conference, TH Darmstadt, March, 1997;
- Ordinal and Symbolic Data Analysis Conference, University of Massachusetts (Amherst), September 1998.

The common theme of the conferences is motivated by the fact that ordinal and symbolic data occur quite frequently, but theoretical tools for handling such data still require further development.

A collection of abstracts of talks from the last meeting appeared as Volume 2 of ENDM, with a Preface by Professor Mel Janowitz, see:

http://www.ulb.ac.be/sciences/ulbmath/osda2000

Further inquiries
They should be addressed by e-mail to the electronic address from which this announcement issues, that is:

osda2000@ulb.ac.be

News from the member societies

News from the IPRCS

Classification and data analysis, as we all know, encompass a wide range of application domains. It is interesting to note how the various member organizations of the IFCS may have been influenced by some quite specific disciplines. IPRCS is no different. IPRCS has been influenced at birth - not uniquely, but fairly clearly nonetheless - by applications related to the processing of visual information, and to the processing of data streams.

This is certainly fertile territory. Areas such as telecoms, science and environmental research, online commerce, and engineering, (to name but a few) are producing continuously petabytes and more of data. To come to terms with modern society - to understand and master the ever bigger tsunamis of data - cluster analysis and data analysis are of critical importance. Traditional data formats such as data arrays, questionnaires, and text, are now being joined by other ever more commonplace data formats - images, video streams, information signals of many types, and all of these stored or arising in a very distributed way. The old problems - "How many clusters are there in a data set?", for example, don't disappear. Instead they reappear in a new light.

(For some further thoughts in this vein on clustering and data analysis over the last four decades, see my Foreword to the special issue of The Computer Journal, Vol. 41, No. 8, 1998, on "Clustering and Classification".)

The idea for IPRCS has grown out of an annual series of conferences, now in its 5th year. This is IMVIP - the Irish Machine Vision and Image Processing Conference, which will be held next at Queen's University of Belfast in August/September 2000. Other annual conferences are also well-established in Ireland, gathering together interesting perspectives from Northern Ireland and the Republic of Ireland at all levels - dissemination of innovative results, letting young researchers get their first slice of the action, and furthering the research/industrial interface.
Such conferences include Applied Statistics in Ireland, Artificial Intelligence and Cognitive Science, and others.

To know more about the IPRCS, it is well worthwhile to bookmark the IPRCS web area at:
http://www.iprcs.org

F. Murtagh, f.murtagh@qub.ac.uk,
http://www.cs.qub.ac.uk/~F.Murtagh

News from the GfKl

The 24th Annual Conference of the German Classification Society took place at the University of Passau on March 15-17, 2000. It was a successful event, with the local organizer Prof. Dr. Gunter Ritter and Prof. Dr. Wolfgang Gaul as the chairman of the scientific program committee. About 200 participants attended the meeting, a remarkable number of them came from Australia, Austria, Belgium, Canada, Denmark, France, Greece, Great Britain, Italy, Japan, Netherlands, Poland, Portugal, Switzerland and USA.
At the University of Passau an interesting and versatile program was presented, with 18 invited plenary and semiplenary lectures as well as more than 100 special talks organized in 30 sections. Special emphasis was placed on the general topic *Classification, Automation and New Media.*
In mathematical methods of *classification* central issues are high-dimensional or other complex data structures. At this conference nonlinearity, dependence structures and non-Gaussianity played a particular role. Special subjects were multidimensional scaling, nonlinear PCA and support vector machines, automatic feature generation, mixture models and the EM-algorithm, projection pursuit, methods of statistical and reinforcement learning as well as applications, in particuar to the classical fields of economics and medicine.

Automation as opposed to manual or brain work is one of the big challenges today. So automatic vs. manual methods of information treatment and their relationships with linguistics had been central topics of this conference. Relating talks dealed with automatic acquisition, discovery, retrieval, creation, transcription, classification, and conversion of data and knowledge. In portfolio management and balancing, e.g., a comprehensive analysis is only possible by automation. The *internet* had a particular appeal as an object of study, here with the application of an old mathematical model, Markov chains, to a brand new problem, that of navigation in the internet.

The third title took account of the *New Media.* There have been presented several talks on the application of fuzzy logic to multimedia applications, the indexing of multimedia documents in the web, object description languages suited for the New Media and applications to diverse domains such as chemistry and the virtual university. Further fields of application presented were:

- Genome Analysis and Systematics
- Medicine and Public Health
- Empirical Economic and Social Research
- Marketing and Market Research
- Finance, Capital Markets and Risk Management
- Linguistics and Information Processing

The Annual Conference 2000 has emphasized the intention of the German Classification Society to form an unifying forum for interdisciplinary work ranging from the foundations to manifold applications of „classification methods“. Presently the proceedings volume is in preparation. We expect the volume to appear at Springer Verlag in spring 2001.

G.Ritter, O.Opitz

News from the JCS

The annual research meeting of JCS was held on February 25, 2000 at the Institute of Statistical Mathematics. In this meeting, Professor Bock took a special talk as an invited speaker. As the contributed papers, the thirteen speakers presented their researches. The abstracts were shown as the following.

(1) Clustering Methods and Generalized Kohonen Networks; Hans-Hermann Bock (Technical University of Aachen, Germany)

The paper establishes a relationship between models and methods well-known from classical cluster analysis and the analysis and construction of Kohonen maps which visualize high-dimensional clouds of data points by the vertices of a (two-dimensional rectangular) lattice. Our approach proceeds by first defining a suitable clustering criterion (*K*-criterion) which measures the fit between the point configuration and the vertex representation, in terms of a weighted sum of Euclidean distances from class centers as in SSQ clustering. Then (variants of) the classical algorithms for constructing or approximating an optimum classification can be applied: the *k-means algorithm*, the *sequential approach by MacQueen*, and the *stochastic approximation approach.*
This criterion-guided approach provides some alternatives to Kohonen's basic 'self-organizing' algorithm and sheds some light on the type of the asymptotically resulting configuration (including topological correctness).
A major advantage of the approach results from the possibility to generalize it to situations where the classes (vertices) are not appropriately represented by class centers (as in a normal distribution case), but where a distribution model with class-specific parameter vectors might be more appropriate and describes better the behaviour of the points inside the same class. This leads to algorithms for 'model-based self-organizing maps' where, e.g., neighbouring vertices represent clusters with similar regression models, principal components spaces etc.

(2) The effect of medical fee revision on lung cancer treatment costs; Ayako Sonoda, Chin Pai Pin, Keiji Yajima (Science University of Tokyo)

The study is a part of the project "Economic evaluation of lung cancer treatments". We examine and study the effect on medical expenses for which the price standards for medicines are reviewed each year. We constructed a cost database grounded on claim forms of one hundred thirty one patients with lung cancer through 1995 to 1998. Since costs for each case depend on the price standards for medicines in the year under review we compare for each medicine two prices based on 1995 and 1998 price standards. The price standards contain various hospital charges, prices for technical skills with regard to various kinds of examinations and tests etc. Although all medical expenses are determined on the basis of price standards, in claim forms there are many aggregated costs not uniquely decomposed to single prices. Simple basic facts and figures will be introduced.

(3) Different Values of Different Age and Sex Cohorts 3: Analyses by Using Hierarchical Clustering and ADCLUS; Akinori Okada (Rikkyo University), Yoshimi Kimura (Osaka University)

Differences among different age and sex cohorts of their values in the Japanese society were studied. Data collected by a nationwide survey conducted in 1995 were utilized in the present study. Responses to 13 items asking about one's values concerning issues in daily life, given by about 2,700 people of age 20 to 70, were analyzed in the present study.
Ten cohorts were formed by combining sex (female or male) and age (20-29, 30-39, 40-49, 50-59, and 60-70 years old).
Product moment correlation coefficients among the 10 cohorts were derived from the standardized responses to the 13 items. The obtained 10 x 10 correlation matrix was analyzed by using the hierarchical cluster analysis and ADCLUS (Shepard & Arabie, 1979).

The obtained results showed two interesting points (a) and (b) below. Firstly, (a) differences among 10 cohorts of one's values seem to be sufficiently accounted for by two factors; one represents differences between extroverted and introverted values, and the other represents differences between materialistic and post-materialistic values (cf. Inglehart, 1990). This is compatible with the results obtained from the analyzes utilizing multidimensional scaling procedures (Kimura & Okada, 1999). Secondly, (b) discrepancies between female and male cohorts of the same age were the largest at their forties, and the smallest at their sixties. Discrepancies increased from twenties through forties, and then decreased from forties through sixties. This is also compatible with Kimura and Okada (1999). The obtained results suggested that discrepancies at forties were distinct from those at other ages. Discrepancies at twenties and thirties varied within extroverted values, and those at fifties and sixties varied within introverted values. On the other hand, discrepancies at forties varied between introverted and extroverted values.

References

[1] Inglehart, R. (1990). Culture shift in advanced industrial society. Princeton, NJ: Princeton University Press.

[2] Kimura, Y., & Okada, A. (1999). Analyzing how to assess one's value among different age and sex groups via multidimensional scaling [summary]. Proceedings of the 23rd Annual Conference of the German Classification Society, p. 89.

[3] Shepard, R. N., & Arabie, P. (1979). Additive clustering: Representation of similarities as combinations of discrete overlapping properties. Psychological Review, 86, 87-123.

(4) Differences between Japan and US in Social Acceptance of "Shikohin"; Shuhei Matsuki (Tobacco Research Institute)

A survey was conducted in both the United States (in 7 northeastern states) and Japan (in Tokyo and 6 other prefectures of the Kanto District) concerning the intake of popular "shikohin", such as coffee, black tea, tobacco and alcoholic beverages. Its purpose was to elucidate the differences and similarities between the two countries in people's awareness in consuming these "shikohin". An examination on the frequency and quantity of intake of each "shikohin", people's awareness of their merits and demerits, and their lifestyles has revealed the following facts.

With regard to the merits of "shikohin," those who consume coffee, black tea, alcohol or tobacco have a greater awareness of their effects than those who do not consume them, concerning almost all questions and for all of these 4 items in both countries. The former highly appreciate in particular such features as 'relaxing' and 'good taste and aroma' in "shikohin".

As for the demerits, or unhealthiness, those who do not consume coffee, alcohol or tobacco have a greater awareness of their effects than those who do in both countries. In addition, Japanese people are less conscious of health hazards of coffee and alcoholic beverages than American people. However, people in both countries are strongly aware of the hazards of tobacco, and there is no difference in this regard between the two countries.

Furthermore, according to the studies on demerits of respective "shikohin" and people's lifestyles, it seems that the intake of alcohol and tobacco is affected by a 'health-oriented trend' in Japan, and, in the U.S., by the same trend as well as a 'religious consciousness.' Regarding the health-oriented trend, Americans tend to endeavor to fulfill goals set for themselves, while Japanese are more resigned to their present health conditions.

(5) Japanese, Japanese Americans, Americans: Data Analysis of Japanese American Surveys; Kazue Yamaoka (Teikyo University), Ryozo Yoshino, Chikio Hayashi (The institute of Statistical Mathematics), Fumi Hayashi (Toyo Eiwa University)

In a comparative analysis of the survey results of seven countries and those for Japanese Americans (JA) living in Hawaii, characteristic features of Japanese-like attitudes were observed in the framework of "Cultural Link Analysis." (Hayashi, et al., 1998) We conducted the Japanese Americans in the West Coast Survey (JAWCS) in 1998 in order to examine whether the enduring Japanese national characteristics have been transmitted to generations of Japanese Americans with Japanese Americans co-researchers (Miyamoto, F., Fugita, S.S., Kashima, T.). The sample were JAs living in King County of Washington and Santa Clara County of California.

Based on the former survey results of the CLA and JAWCS results, responses of the questions on interpersonal relationships, sociological attitudes and behaviors, etc. were analyzed.

Special attention was given to a topic used to specify the clusters among Japanese (J), JA in Hawaii (JA-HA), JAWCS, and Americans (A). In the analysis based on the questions related to Japanese characteristics, a stereotype relationship (J)-(JA-HA, JAWCS)-(A) was observed.

On the other hand, the clusters based on sociological attitudes and behaviors affected by group specific circumstances. The study was supported by a Grant-in-Aid for Scientific Research A(2) (No. 10308007) from the Japan Ministry of Education, Science, Sports, and Culture, awarded to Ryozo Yoshino.

References
[1] Hayashi,C., et al.: "Kokumin-sei 7 kakoku hikaku" (in Japanese) Comparative study of National character in seven country study. IDEMITU SHOTEN, 1998.

(6) Comparative Study on National Character and Data Mining - Evaluation of Data Quality and Data Analysis; Chikio Hayashi (The Institute of Statistical Mathematics)

This presentation discusses two problems in cross-societal sample surveys. The first concerns the translation of questionnaires. Suppose that a question written in Language A is translated into Language B. Then the question in Language B is translated back into Language A. The original question is then compared with the "back translation." To further assure the questions' similarity of meaning, a split-half survey is conducted; one-half receiving the original question and the other half the back-translated question, thereby investigating the questions' comparability. The second problem concerns the survey methodology, particularly the sampling and data collection methods used by survey companies. If strict random sampling is employed, few serious problems are likely to arise, apart from possible problems relating to the data collection method and procedure, non-response and response errors. In many countries, however, survey companies use only quota sampling, which gives rise to many technical problems in sampling design and data collection. Survey companies have their own particular skills and in many cases their techniques and know-how are considered proprietary and are therefore kept confidential. But data quality depends upon these techniques and know-how. Here, some examples of these phenomena and problems are described, focusing in particular on the differences one can observe in the results.

(7) Fuzzy Clustering with Crisp Regions; Norio Watanabe (Chuo University), Tadashi Imaizumi (Tama University)

A new fuzzy clustering algorithm is proposed by introducing crisp regions of clusters. Results of the usual fuzzy k-means are fuzzy and only the data on the center can have the membership value one. On the other hand the new method produces the crisp regions where membership values are given by one or zero. The area between crisp regions is a fuzzy region. The crisp regions in this article are defined as spheres whose centers are the centers of gravity. When the radii of spheres are zero, the proposed method is identical to the usual fuzzy k-means. Numerical examples are demonstrated.

(8) Relational Fuzzy c-Means for 3-way Data; Mika Sato-Ilic (University of Tsukuba)

Relational fuzzy c-means method (RFCM) was extended for applying 3-way data which consists of dissimilarity over several times. Two extended methods were proposed.
One method used multi-functional criteria with the weighting method. This became essentially the same as a method which combines the dissimilarities over times using the weighting method, that is described by the functions of the solutions of the evaluation function of RFCM and its algorithm. In this case, we can get the unique clustering result through the times.
The second method used the idea of super-matrix and gets clustering results for each time. These clustering results are comparable under the same clusters.
With this method, in the super-matrix, we define the changing situations using the non-diagonal matrixes of the super-matrix whose elements are defined by values of a function which is defined as the changing situation of dissimilarity of a time to another time.
In order to compare the results between the two methods, an evaluation function which measures the homogeneity situation of clusters is proposed. Formerly, the evaluation of the results of fuzzy clustering, that is degree of belongingness for clusters, has used the partition coefficient. However, in the above cases, we cannot use this evaluation function, because of the difference of the data between the times. So, the idea of homogeneity of clusters is introduced base on homogeneity analysis.
Several numerical examples were demonstrated to show the performance of the proposed issues.

(9) Fuzzy Clustering for Time-Dependent Similarity; Yoshiharu Sato (Hokkaido University)

The aim of this paper is to recognize a change of the latent structure of a similarity data through a change of clusters.
However we can get the clustering result for the similarity at each time point, it is difficult to compare the results between different time points, even if the number of clusters are identical. To do this, we move continuously the similarity values between two adjacent time points using a convex combination and apply the additive fuzzy clustering model. Then, tracing the clustering results, we could compare the results between two different time points.

(10) Projection and classification by minimizing loss function; Takahiro Nakamura (Graduate University for Advanced Studies), Yasumasa Baba, Noboru Ohsumi (The Institute of Mathematical Statistics)

Homogeneity analysis is a method for describing data structure by minimizing loss functions which define homogeneity. Various loss functions that express homogeneity have been defined.

Nakamura expands the homogeneity and proposed a method for describing data structure, PML (projection method by minimizing loss function).
This method projects variables onto lower dimensional space and describes data structure in the space.
Nakamura proposed a method of clustering variables which utilizes a loss function. The method uses loss functions as criteria of clustering variables. Minimizing the loss function is the same as a projection of variables for certain spaces. The loss obtained here is the loss of projection. This method searches partitions with small loss, spaces which make a good visualization of the partitions and plots of variables onto the space by minimizing the loss function. We use differences of profiles of variables for clustering and visualization of data structure. This method was applied for physical fitness data (Kishida, Shindo, Ohsumi). This data was used to estimate human physical fitness. The principal component analysis and clustering method by minimizing loss function was applied to 15 items of the data. These items were classified and projected onto a space. We compare the scatter diagram of factor loading to the scatter diagram of variables in the space by the clustering method.

(11) How to extract information from call-takers' reports; Masayuki Morohashi and Tadashi Imaizumi (Tama University)

Out goal is to establish a way of applying data mining techniques to narratively described text data such as call-takers' reports. In this paper, we discussed how to extract information from such texts.
We focus on the availability of using natural language processing techniques from the points of what kinds of information they can get, and of how precise the results are. The key of those discussions is how to construct the application-oriented dictionary (includes pairs of word and its category), which determines the capability of analyzing texts.

(12) An Exploration of Organization Genome - Organization Chart Encoding Approach -; Takao Ishizuka(ASIA University)

We suppose "organization genome" as a concept corresponding to the inherited or self-duplicated specification of corporate organizations. The organization chart is one of the static specifications of organizations and doesn't represent the whole information of the organization. However, the organization chart often decides its behavior and is reflected in the organization culture. Accordingly, organization chart data can be regarded as organization genome.
This article presents an encoding method and some data analysis of organization charts.
Generally, organization charts represent the lines of command and can be described by the relationship between a sub organization or unit and its upper organization. Graphical and hierarchical organization charts can be encoded into the correspondence table between units, and the table can be transformed into an adjacency matrix.
We develop some data analysis methods of the encoded organization charts. Distance matrix between units and the concerned statistics serve very useful summary information on the organization chart. The value of the total or mean distance between units is an index of the complexity of the organization.
Many problems are remained. First, the discrepancy between the actual organization and its chart is often observed. For example, one unit leader currently holds another unit posts. Consequently, actual distance between two units may be zero in spite of positive value computed from the organization chart. The actual state of one-person-multi-position supposes another organization genome. Secondly, there are some problems on encoding the organization charts. For example, how to cope with staff organizations is one of the next problems. Thirdly, a method of comparison and analysis of multiple enterprises organization charts must be established.
Fourthly, the relationship between the organization performance data, i.e. financial data and organization genome must be analyzed. Finally, we must explore what the organization genome of excellent companies should be.

(13) Attitudes towards Educational Evaluation of University Students in Department of Education; Osamu Yoshimura (Okayama University)

Attitudes towards educational evaluation of university students were investigated. By analyzing free answers to a question "What is educational evaluations for you?", views of educational evaluation were classified into following five types: "being judged their whole value", "begin measured their traits", "being appreciated their efforts", "being given an diagnosis of their characteristics", "being under the control of others".
It was found that about 70% of the students who wanted to be a teacher understood educational evaluation to judge one's whole value and over 60% of the students who wanted not to be a teacher felt to be under the control of others. Moreover, global self-esteem of the students who have the view "being judged their whole value" was lower than other students.

(14) Analyzing Textual Data on Fatigue Awareness, Frustration and Aggressive Attitudes; Kiyoharu Doi (Miyazaki Sangyo-keiei University), Noboru Ohsumi (The Institute of Statistical Mathematics)

Open-ended questionnaires and the four kinds of usual closed-type questionnaires about fatigue awareness are investigated for the students of a university in Japan to compare the relationship between their free answers and their choices of the questionnaire items. Then, the textual data from their free answers and the selected choices are simultaneously analyzed and evaluated using InfoMiner that is a specific software for analyzing textual data and the choices in the usual closed questionnaires such as multiple response type. As the results of statistical analyses, based on the extracted words, we could find out three groups concerning with the fatigue awareness, frustration and aggressive attitude.
Respondents in the one of the groups are exhibited quite fatigue awareness in their free answers and highly scored on the results of responses of the closed questionnaires for asking the fatigue attitudes. Furthermore, with respect to frustration and aggressive attitude, we could examine several similar points among them. Thus, in concluding, the difference and similarity of the measurements between textual data from free answers and the results of usual closed questionnaires are clearly examined.

News from the VOC

The Dutch and Flemish Society for Ordination and Classification (Vereniging voor Ordinatie en Classificatie, VOC) was founded in 1989. To celebrate the 10^{th} anniversary, a two-day conference was organized last November. The theme was: "Everything on time", ordination and classification with time-bound data. The conference program was divided into three parts: 1) reference curves and related problems, 2) reconstruction of time sequences, 3) miscellaneous modelling problems in time series data. Members and non-members of the VOC presented research and applications. There were two key-note speakers of fame: Tim Cole (reference curves) and Jim Ramsay (functional data analysis).
Fifty people attended the conference, which was held in the historic buildings of the Rolduc conference center (an old abbey) in Kerkrade. Abstracts of the presentations and information about the speakers can be found at our home page (www.voc.ac).

The VOC is a healthy society. The membership rose last year by over 10% to nearly 120. The meetings (two each year) are well attended.

The next meeting will be at April 7, in Groningen.
The subject is "Classification with Mixture Models".

Details can be found on our website:

www.voc.ac

News from the SFC

(1) Results of the SFC elections

- Pascale KUNTZ has been reelected "Editor of the SFC Bulletin" (1999-2002)
- Andre HARDY has been reelected "Secretary of the SFC" (1999-2002)

(2) The general assembly of the SFC will be held during the "Seventh Conference of the International Federation of Classification Societies" (Namur, Belgium, July 11-14, 2000).

(3) The issue number 12 of the "SFC Bulletin" was published in October 1999.

(4) The next SFC meeting will be held in Pointe-a-Pitre (Guadeloupe) in 2001. Contact: sfc2001@univ-ag.fr
Information available in this newsletter.

(5) IFCS-2000 (Namur, July 11-14, 2000) - Hotel Reservation

If you did not book your hotel yet, it is quite urgent to do it soon. The IFCS-2000 local committee encourages you to book a room at the Novotel Hotel.

(Fax: +32-81-461990, Price (breakfast included): Single and Double Rooms: 2500 BEF (62.5 EURO)).

The Novotel Hotel is located beside the Meuse river, and will be linked with the IFCS-2000 meeting by a twice-a-day shuttle, plus an all-day free shuttle service on request.
Moreover, there is a city bus (Bus n°4, Quai D in front of the Railway Station) to/from the hotel every 30 minutes.

All the information concerning IFCS-2000 is available on the website:

http://www.fundp.ac.be/~ifcs2000

Andre HARDY
SFC Secretary
IFCS-2000 - Secretary Local Organizing Committee
E-mail: Andre.Hardy@fundp.ac.be

Meetings of member societies

GfKl meeting
25th Annual Conference – University of Munich – March 14-16, 2001

The German Classification Society (GfKl), a member of the IFCS (International Federation of Classification Societies), will hold its 25 th Annual Conference at the Ludwig-Maximilians-University of Munich. The scientific program will include plenary lectures and lectures in special sections. Interdisciplinary aspects and interrelations with respect to theory and application will gain essential priority. Emphasis will be placed on **Explorative Data Analysis in Empirical Research.**
Typical **topics of interest** and relevant **fields of application** are:

- Mathematical and statistical methods for data analysis and classification
- Econometrics and empirical economic research
- Representation and visualization of data
- Computer-aided methods of statistics
- Bayesian data analysis
- Knowledge-based methods and classification
- Data mining: IT-driven aspects
- Data mining: methods and applications
- Web mining
- Information systems
- Marketing and retailing
- Market research
- Finance, capital markets, and risk management
- Production, operations, and controlling
- Concept analysis and conceptual classification
- Environmental sciences, climatology, and geography
- Nutritional and health sciences
- Biology, biometrics, and chemistry
- Genome analysis
- Medicine and neuro-physiology
- Linguistics and information retrieval
- Subject indexing and librarianship
- Archaeology

Within the framework of the conference, workshops, tutorials, and software demonstrations are scheduled.
Conference languages are German and English. A Proceedings volume will be published in English by Springer-Verlag. Interested persons will obtain further information if they send a preliminary registration.
Participants who intend to present a talk are requested to submit an abstract by **November 15 th , 2000** to the local organizer Prof. Dr. M. Schwaiger.
Accepted abstracts will be included in a summary volume that will be distributed to conference participants. Speakers will be notified of acceptance by 15 January 2001.

Upon acceptance, a complete manuscript that will have to pass a review process for inclusion in the proceedings volume can be submitted. Formatting instructions for abstracts and manuscripts can be obtained at:

http://www.efoplan.bwl.uni-muenchen.de/gfkl.html

Preliminary program committee
M. Schwaiger (Munich), F. Critchley (Birmingham, UK), R. Decker (Bielefeld), L. Fahrmeir (Munich), W. Gaul (Karlsruhe), B. Lausen (London, UK), Y. LeChevallier (INRIA, Le Chesnay Cedex, FR), K. Obermayer (TU Berlin), O. Opitz (Augsburg), G. Ritter (Passau), A. Sokolowski (Krakau, PL).

Local organizer
M. Schwaiger

Correspondence to:
Prof. Dr. M. Schwaiger, EFOplan, LMU Munich,
Faculty for Business Administration, Kaulbachstr. 45, D-80539 Munich
Tel.: ++49 89 / 21 80 – 56 40, Fax: ++49 89 / 2180 – 56 51, E-mail: gfkl2001@bwl.uni-muenchen.de

Bank account
M. Schwaiger/GfKl-Tagung 2001, Account-No. 022 45 43 01, Deutsche Bank 24 Munich,
bank code no. 700 700 24

If you intend to participate in the 25th annual conference of the German Classification Society (GfKl), you can write to:

Prof. Dr. Manfred Schwaiger
Seminar für Empirische Forschung und Quantitative Unternehmensplanung
Universität München
Kaulbachstr. 45
D-80539 München

CSNA Meeting
Annual meeting, Montreal, June 8-11, 2000

Notice: The Classification Society of North America has been compelled to move its annual meeting one week earlier than originally planned. The new schedule is for the shortcourses to occur on June 8, followed by our reception that evening. The regular conference will begin on the morning of June 9, and last until noon on June 11. The CSNA banquet will take place on Saturday night, June 10. We apologize for any inconvenience this change may cause.
More information can be found at:

http://www.crt.umontreal.ca/GERAD/CSNA/index-en.php3

CLAD Meeting
JOCLAD 2000, Lisbon, March 23-25, 2000

JOCLAD-2000, the 7th annual meeting of the CLAD, will be held next March 23-25 in Lisbon. The Institute for International Scientific and Technological Co-operation (ICCTI), the Foundation for Science and Technology (FCT), the National Institute of Statistics (INE), the French Embassy and the Luso-American Foundation for Development (FLAD), have already given their support to the meeting.

The JOCLAD 2000 will take place at the Institute for Development of Enterprise Management, INDEG. Official languages are Portuguese, French and English, as usual.
Provisional program includes:

Invited Sessions – by Peter Bryant (Univ Colorado/Denver, USA), *Learning multivariate data analysis by case studies: Some issues and examples;* Michel Lejeune (Univ. Grenoble 2, France), *Multivariate analysis for large amounts of missing data;* Annie Morin (IRISA, Univ de Rennes, França), *Utilisation d'analyses factorielles des correspondances successives à partir d'un tableau de données, pour des objectifs différents;* Leon Bobrowski (Institute of Biocybernetics and Biomedical Engineering PAS, Warsaw, Poland), *Data mining: Selected methods and applications*

Thematic Sessions
Teaching / Learning multivariate data analysis.
Multivariate analysis with missing data.
Application of data analysis in environmental sciences.
Analyse des données textuelles. Réseaux neuronaux pour l'analyse des données textuelles. "Validation de Structures de Classification" and SODAS Programmes.

Software Sessions - by SPSS (*Methods and Software. Applications*) and SAS Portugal
(*Knowledge creation. Data mining to obtain knowledge)*

Software Course on SPAD N + SPAD T - by Alain Morineau (CISIA, França):
Methods and Software: Applications
Panorama des méthodes et du software
Expériences de data mining et d'analyses des données avec SPAD
Le rôle de la statistique dans le data mining.
Exemples d'applications: segmentation, scoring, classification

More information can be found at the CLAD web page.

e-mail: cladlead@fc.ul.pt
clad.lead@mail.telepac.pt
web page: www.fpce.ul.pt/~cladlead/

Helena Bacelar-Nicolau, President of CLAD
hbacelar@fc.ul.pt

SFC Meeting,
8th conference, Guadeloupe, November 2001

After the 7th Conference of the IFCS (International Federation of Classification Societies) which will be held in July in Belgium, the annual rhythm of the SFC (Société Francophone de Classification) meeting will resume with its 8th Conference.
It will be organised in Guadeloupe at the French West Indies and Guyana University in November 2001.

It will be jointly organised by the TRIVIA computer scientist team and the statistician group INSTAN. Because of the exceptional trip duration, the Conference will be held on a whole week, two days for a special course, and three days for the standard sessions.
The "Simon Régnier" award will be discerned to a young researcher for his applied and/or theoretical results on a special Caribbean data base.

The local organisation committee will do his best for this above all scientific meeting to be pleasant!

Conferences of IFCS member societies (see IFCS website):

- CLAD 2000, March 23-25, 2000, Lisbon
- CSNA 2000, June 8-11, 2000, Montreal
- GfKl-2000, March, 15-17, 2000, Passau
- GfKl-2001, March, 14-16, 2001, Munich
- IPRCS (IMVIP 2000), Aug 31- Sept 2, Belfast
- SFC-2001, November, Guadeloupe
- SIS 2000, 26-28 April 2000, Florence
- VOC spring meeting , April 7, 2000, Groningen

International Congress on INTELLIGENT SYSTEMS AND APPLICATION
University of Wollongong, (near Sydney), Australia, December 12-15, 2000

The ISA'2000 Congress aims to provide researchers and practitioners from academia and industry with a forum to report on the latest developments in intelligent systems and their applications within the four major areas of computational intelligence, interactive and collaborative computing, industrial systems, and biologically inspired systems. The Congress will also provide a unique opportunity for dialogue and synergy between scientists and engineers from different backgrounds with the common interest in intelligent systems.
The congress consists of the following four Symposia:

1. *Computational intelligence (CI'2000)*
   Chair: Prof. Franz J. Kurfess, Canada
   Email: franz@cs.concordia.ca
   for details see http://www.icsc.ab.ca/151-info.htm

2. *Interactive and collaborative computing (ICC'2000)*
   Chairs: Prof. Yoneo Yano and Prof. Hiroaki Ogata, Japan
   Email: yano@is.tokushima-u.ac.jp
   ogata@is.tokushima-u.ac.jp
   for details see http://www.icsc.ab.ca/152-info.htm

3. *Industrial systems (IS'2000)*
   Chair: Prof. Edward Szczerbicki, Australia
   Email: mees@cc.newcastle.edu.au
   for details see http://www.icsc.ab.ca/153-info.htm

4. *Biologically inspired systems (BIS'2000)*
   Chair: Hans-Heinrich Bothe, Denmark
   Email: hhb@it.dtu.dk
   for details see http://www.icsc.ab.ca/154-info.htm

Submission of papers
Prospective authors are requested to either send a draft paper (maximum 7 pages) or an extended abstract for review by the International Program Committee. All submissions must be written in English, starting with a succinct statement of the problem, the results achieved, their significance and a comparison with previous work.

Submissions must be received by March 31, 2000.
Submission by electronic mail is strongly recommended to:

operating@icsc.ab.ca

Important dates
March 31, 2000: Submission deadline
May 31, 2000: Notification of acceptance
July 30, 2000: Delivery of full papers
December 12-15, 2000: ISA'2000 congress

Congress general chair
Prof. Fazel Naghdy
School of Electrical, Computer, and Telecommunications Engineering
University of Wollongong
NSW 2522 / Australia
Email: f.naghdy@uow.edu.au

Congress organizer
ICSC International Computer Science Conventions
P.O. Box 279
Millet, AB T0C 1Z0 / Canada
Fax: +1-780-387-4329
Phone: +1-780-387-3546
Email Operating Division: operating@icsc.ab.ca
Email Planning Division: planning@icsc.ab.ca

http://www.icsc.ab.ca/isa2000.htm

IFCS Newsletters on Internet:
http://edfu.lis.uiuc.edu/~class/ifcs/newsletter.html
The newsletters are available as PDF files, to be read with AcrobatReader. Click on AcrobatReader to download.

IFCS Homepage:
http://edfu.lis.uiuc.edu/~class/ifcs/
Designed and maintained by David Dubin.
The website now contains:
information on the Namur conference, on travel awards for young researchers, the IFCS Constitution and By-Laws, the IFCS newsletters, and pointers to the websites of the member societies.

News for the coming issue of the IFCS can be sent to:

Paul.Deboeck@psy.kuleuven.ac.be

Text files are by preference in ascii, wordperfect, or word, with pc format (mac is difficult, although not impossible). For graphical materials, by preference BMP, GIF, or PCX are used.

Editor:
Paul De Boeck, Department of Psychology,
University of Leuven,
Tiensestraat 102, B-3000 Leuven,
Belgium.
tel: +32 16 325980 fax: +32 16 215916