International Federation of Classification Societies Newsletter

Number 27   **May 2004**

President: Henk Kiers   Editor: Krzysztof Jajuga

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

Associaçáo Portuguesa de Classificaçào e Análise de Dados, British Classification Society, Central American and Carribean Society of Classification and Data Analysis, Classification Society of North America, Gesellschaft für Klassifikation, Irish Pattern Recognition and Classification Society, 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 biennial conference, and supports the Journal of Classification.

Contents

I would like to inform that IFCS Council completed voting on IFCS Logo, which appears on upper left corner of the Newsletter. Also IFCS Council approved new Constitution, By-laws and voting procedure

Krzysztof Jajuga, Publication Officer

I would like to wish the participants and the organizers of IFCS-2004 Chicago conference all the best during the conference.

Krzysztof Jajuga, Publication Officer

IFCS Homepage:

http://www.classification-society.org/
Designed and maintained by David Dubin.
The website contains among others:
the IFCS Constitution and By-Laws, the IFCS newsletters, and pointers to the websites of the member societies.

A WORD FROM THE PRESIDENT

On January 1, I had the honour and pleasure to succeed Carlo Lauro as IFCS president. As we had had frequent contacts the year before, and Carlo had documented his initiatives very well, the transition went smoothly, for which I wish to thank Carlo wholeheartedly. The first four months I have mainly worked on the finalization of the initiatives by my predecessor. This has now led to various concrete results.

First of all, as you will have noticed on the front page of this newsletter or on our website, we have a

new logo! It has been prepared by Handan Arik, and can be seen to express the network structure of the IFCS, but also classification as a data analysis method is visualized by means of the tree structure in the logo. In addition to the logo, she has also developed IFCS official stationery, and as an extra has offered to redesign the IFCS website accordingly, free of charge, for which we are very grateful.

Secondly, the significant revision of the IFCS Constitution and By-laws by Hans Bock and Jean-Paul Rasson's Committee of Sages is currently being voted on. Besides the elimination of various obsolete articles, some important changes have been proposed on the structure of the Council and the terms of officers, in order to keep the continuously growing federation manageable.

Next, new Finance Committee members (Javier Trejos, Yoshiharu Sato and Iven van Mechelen) and a new chair of this important committee (Iven van Mechelen) have been elected. The new committee is working effectively on the previous as well as current financial reports, thereby making up for the inadvertent delay that had been incurred in the past.

Furthermore, of course, the preparation of our IFCS2004 meeting in Chicago is in full swing. David Banks and Buck McMorris have ensured that the proceedings volume has been sent to the publisher, and they are currently working on the conference program, with more than 150 talks, including an excitingly large number of invited talks and sessions. Related to this, the Travel Awards Committee (chaired by Andrzej Sokolowski) has recently selected and informed its winners, so that again the attendance by young promising researchers has been given a further stimulus, thanks to the fund that was so kindly established by the late Chikio Hayashi. Although a new donation to this fund was received, further means will be sought for maintaining this fund.

Finally, important contacts are being maintained by Carlo Lauro with classification groups from Greece, South Africa and China. Further steps have been taken in the past months, and I hope that we can welcome these groups as new classification societies in the IFCS before long.

Having now harvested the fruits of my predecessor, the time is there to work on new initiatives. My primary concerns for the near future are:

1. Stimulating direct participation of individuals in IFCS matters.
2. Education and usability of classification and data analysis.

As to the first issue, it is my aim to fruitfully use the by now almost world-wide availability of the internet and email. This allows for undelayed accessibility of IFCS information to individuals. Announcements, newsletters but possibly also substantive discussion matters on classification and data analysis can be sent out effectively by email directly, rather than through the societies. As an added benefit, interested individuals in regions that do not (yet) have a classification society can be kept informed on matters related to classification and data analysis. Of course, in all this, cautious action is required to avoid public availability of email addresses.

As to the second issue, in my opinion, the society should foremost be a forum on what has rightly been dubbed "Data Science", dealing with all aspects of collecting and analyzing data. This not only covers classical mainstream statistics, but also statistically informal approaches such as Cluster Analysis, Multidimensional Scaling and Data Mining. The IFCS is an outstanding platform for dealing with data analysis techniques that are traditionally ignored in mainstream statistics. With its broad view on Data Science, all approaches to data analysis are welcome, and hence new approaches and classic approaches can be confronted but also synthesized, in an effort to create the best of both worlds. The IFCS could and should play a leading role in these new developments in Data Science. An important aspect here is that IFCS should foster the development of techniques with a keen mind on the applications for which the techniques are mentioned. In particular, techniques must be made usable for the researcher who collects the data (not only for the statistician or classification expert). Therefore, techniques should be user-friendly and transparent, ensuring that the researcher who analyzes the data keeps a clear idea on what the analysis does and what the results display, and the interpretation of results should be crystal clear. This implies, last but by all means not least, that IFCS should pay attention to how to teach our techniques to students and other prospective users. Thus, IFCS should promote an integrative view on developments in research and teaching of Data Science. One step in this direction is the round table session on this topic, organized by Helena Bacelar at IFCS2004. Other steps will be the organization of further IFCS sponsored courses on classification and data analysis.

Above, I have informed you about recent activities and results, as well as my future aims. Any suggestions from you are extremely welcome, because, after all, the fruitfulness of the IFCS depends on the participation of its members. Let me close by expressing my hope to see many of you in Chicago.

Henk Kiers

IFCS President
University of Groningen, The Netherlands
email: h.a.l.kiers@ppsw.rug.nl

OUTGOING PRESIDENT LETTER

Dear Council members, Dear Committee members,
Dearest IFCS Friends,

the end of 2003 coincides with the end of my term as
IFCS President. Thus, it is time to make a balance of
the past with a look at the future of the Federation. It
seems like yesterday when on July 17th, 2002, I gave
my incoming President statement to the Council
meeting in Krakow. It comprised a list of challenges
and goals that I wished to achieve together with you
during my Presidency.
There were actually three main proposed objectives:
make the life in the Federation more participative,
expand the influence of the Federation, develop an
effective publication policy. These objectives were
then joined by the revision of the Statute and By-
Laws aiming at setting up a more effective reference
framework for the achievement of the mentioned
goals and for coping with the needs and new
challenges to be faced by the Federation in the Third
Millennium.
Beside the traditional committees (Finance
Committee, Long Range Planning of Conferences,
TAP, Nomination Committee), several new ones
have been established in order to prepare and support
the achievement of the proposed objectives:
- the *Communication Committee,* chaired by
  David Dubin, with the aim of improving the
  communications within IFCS and of making
  recommendations about the IFCS website;
- the *Education Committee*, chaired by Helena
  Bacelar-Nicolau. for exploring new ways in
  which IFCS can better disseminate the
  science of classification and related works;
- the *Publications Committee,* co-chaired by
  Krzysztof Jajuga and Paul De Boeck, for
  making recommendations on the future IFCS
  publication policy, namely regarding
  journals and monographs;
- a *Committee of Sages*, co-chaired by Hans
  Bock and Jean Paul Rasson, consisting of
  past IFCS presidents, was established with
  the aim of reviewing the Constitution and
  By-Laws.
These Committees started their work in December
2002. A detailed report of their activities is going to
be published in the issue of IFCS Newsletter in 2004.
At this time, I like to remind some of the results that
have been achieved during this period with respect to
the goals being set in my statement.
As to the first objective, in line with a new
participative model in the institutional activities of
the Federation, the membership in the committees
has been opened to qualified representatives from
each society of IFCS that did not already belong to
the Council.

A prominent role has been assigned to past-
presidents for the revision of the Statute. The new
Web site, whose preparation is in progress, already
functioning for a new electronic voting system, will
allow to improve the communications within IFCS
thus making possible forums and news groups, E-
newsletter, dissemination of articles, software and
membership directories.
A special care has been taken of the young people by
setting up a programme of summer schools that has
begun with the IFCS Summer School in Lisbon. The
organisational model of this School, particularly
appreciated by its participants, will constitute a very
useful reference for the forthcoming initiatives.
The transversal working groups on specific themes
represent, nowadays, an important reality of
scientific exchange within the Federation. Beside the
group on Symbolic Data Analysis (SDA) that is
already running, the Correspondence Analysis and
Related Methods Network (CARMEN) was
established in the past days with more than 90
participants.
As to the second objective, the relationships and
liaisons with international scientific associations
(such as IASS, IASC and ISI) have been intensified
in the past two years. Results of these actions include
the summer school in Lisbon (in co-operation with
IASC and IASS), the IFCS presence in the scientific
programmes of Compstat 2004 in Prague, ISI 2003 in
Berlin and ISI 2005 in Sydney. IFCS has also opened
to new scientific communities such as the PLS one
by co-sponsoring the PLS meeting in Lisbon
(September 2003) where useful stimuli were drawn
for integrating new algorithms in classification
problems. IFCS also enjoyed the interesting
participation in the International Meeting of the
Psychometric Society (Chia Laguna, July 2003) by
means of organising some specialised sessions.
Finally, it is honourable for all of us to mention the
interest shown by associations and classification
groups from Greece, South Africa and China in
joining IFCS. We all truly wish that these
perspectives may soon become reality.
All these activities have led to a greater visibility of
the research on the themes of classification and data
analysis as well as to a major international credit of
our Federation.
We have also worked, in co-operation with Handan
Arik on the drawing of a new logo for IFCS. This
logo should appear on any material, document,
leaflet, etc. related to IFCS and thus serve as a
trademark that better identifies our Federation and
makes it more easily recognisable by the
international scientific community.
The exploratory work of the Publication Committee
has allowed to screen possible alternatives of a direct
engagement of the Federation in editorial activities.
Given the importance of this topic that has

systematically come into discussion since many years, the Committee has not yet reached a definite proposal to submit to the Council meeting in Chicago. It is desirable that, in order to prepare a document with recommendations, all Council members let Krzysztof and Paul have their own comments on the Committee's report.

The work done by the Committee of Sages led by Hans and Jean Paul, has finally yielded a precise and detailed revision of the Statute and By-Laws. Henk Kiers and myself have continuously been in touch with this committee during the preparation of these documents also by means of some proposals and discussions. Therefore, we fully agree on their contents. I have the great pleasure of attaching to this last e-mail of mine as President, the proposal by the Committee of Sages.

As a whole, several results have been achieved while other actions are going to yield their effects in the next future. Several other issues have been faced but they still require further widening and discussion by the Council. For sure, the job done has set the basis for a new future of IFCS.

For me, the last two years have been filled by intense and passionate work. Henk is now going to take over. I have shared several projects for IFCS with him. I am sure that he will be able to do a lot for the progress of the Federation especially by feeding IFCS with new ideas.

I wish Henk all the best for a nice job and I place my help and experience at his complete disposal. Moreover, I ask all of you to provide him with the same undertaking and the same valid co-operation that you offered me in the past two years.

In concluding my office, I sincerely wish to thank all of you and express my gratitude for the effective work you have done. It was an essential flavour for the achievement of the mentioned results. Special thanks go to Jean Paul for introducing me to the most relevant issues for IFCS and for supporting me in the difficulties encountered during my term, and to David who has meant to me, and to IFCS as a whole, an irreplaceable reference.

I take this opportunity to send you all and your families my most heartfelt wishes for a 2004 full of peace and contentment in the sign of friendship.
See you in Chicago for a successful IFCS'04!

Carlo Lauro

UPDATE ON IFCS-2004, JULY 15-18 IN CHICAGO

The arrangements for the 2004 meeting of the International Federation of Classification Societies are essentially in place. The conference website (at www.classification-society.org/ifcs04/) has a copy of the invited program, automatic on-line registration, and information about housing options, as well as other useful information. The copies of the refereed proceedings are being printed up this week, the lecture halls and other conference facilities have been reserved, and the social arrangements are all confirmed.

We expect about 300 attendees, and besides the scientific program there will be a cruise on Lake Michigan, a traditional Chicago cookout, and the conference will close with a whiskey tasting (we encourage people to stay over Sunday night, and try to sleep in on Monday morning).

However good the social arrangements, the real reason to come is the research. We are fortunate to have Lynne Billard, Gilles Celeux, Pedro Domingos, Anuska Ferligoj, Wolfgang Gaul, David Hand, Bruno Leclerc, Taerim Lee, Regina Liu, Masahiro Mizuta, Fionn Murtagh, Carey Preibe, Jim Ramsay, Michael Steel Vladimir Vapnik, and Rosanna Verde as invited speakers. And there are many other distinguished researchers participating in the program – we are particularly impressed by the many excellent contributions we have received from new researchers.

So we urge everyone to move swiftly to register and to make their travel arrangements. It will be a great meeting, and we look forward to seeing as many of the IFCS members as possible. We hope we that everyone will enjoy the Chicago meetings as much as we did the meeting in Cracow!

David Banks
banks@stat.duke.edu

IFCS SECRETARY'S REPORT

The IFCS is working on several important fronts. First, I am happy to report that there is a major ballot in progress, regarding whether or not to adopt several important changes to our Constitution and By-Laws that reflect recommendations made by the Committee of Sages that Carlo Lauro appointed to consider strategies for modernizing our arrangements. The balloting closes on May 16, and the results will be announced soon after.

Also, I am pleased to report that the Finance Committee is working hard to prepare a report on the records from our Treasurer, Taerim Lee. The Finance Committee consists of Iven Van Mechelen (chair), with Rosanna Verde, Javier Trejos, and Yoshiharu Sato. We are confident that their report will be ready in time for the IFCS Council meeting in Chicago, and that the report will bring us back onto the timetable for review that was originally intended.

Additionally, the IFCS Council has approved a new logo, which appears in this newsletter. The logo was designed by Handan Arik and we certainly hope that everyone likes it. The Executive Committee is deeply appreciative of the outstanding work that Handan did to help improve our public image.

Finally, to look ahead a little, let me note that the chairs of various committees should start to think about preparing their reports for the Council meeting, and the members of the member societies should contact their representatives if there is business that should be proposed to the Council.

Also, we have had one offer, from Anuska Ferligoj and Vladimir Batagelj, to host the 2006 IFCS meeting in Slovenia. If there are any other members who would like to host, we need to hear from them very soon – the final decision will be made at the Council meeting in Chicago.

David Banks
IFCS Secretary

NEWS FROM GfKl

At the last GfKl annual conference (Dortmund, March 9-11, 2004) a new board has been elected for the period of 2004 - 2007.

President:
Prof. Dr. Claus Weihs, Lehrstuhl für Computergestützte Statistik, Fachbereich Statistik, Universität Dortmund, Vogelpothsweg 78, D-44221 Dortmund, Tel.: +231/755-4363, Fax: +231/755-4387, e-mail: weihs@statistik.uni-dortmund.de

1st Vice-president:
Prof. Dr. Reinhold Decker, Lehrstuhl für BWL und Marketing, Fakultät für Wirtschaftswissenschaften, Universität Bielefeld, Postfach 100 131, D-33501 Bielefeld, Tel.: +521/106-3936, Fax: +521/106-2994, e-mail: rdecker@wiwi.uni-bielefeld.de

2nd Vice-president:
PD Dr. Berthold Lausen, Institut für Medizininformatik, Biometrie und Epidemiologie, Universität Erlangen-Nürnberg, Waldstr. 6, D-91054 Erlangen, Tel.: +9131/85-25739, Fax: +9131/85-25740, e-mail: Berthold.Lausen@rzmail.uni-erlangen.de

Treasurer:
Prof. Dr. Hermann Locarek-Junge, Lehrstuhl für Betriebswirtschaftslehre, insbesondere Finanzwirtschaft, Technische Universität Dresden, Mommsenstr. 13, D-01062 Dresden, Tel.: +351/463-5572, Fax: +351/463-5404, e-mail: locarekj@finance.wiwi.tu-dresden.de

Other members (in alphabetical order):
Prof. Dr. Werner Esswein (Dresden), Prof. Dr. Andreas Geyer-Schulz (Karlsruhe), Prof. Dr. Gunter Ritter (Passau).

Further elections have been conducted in the Working Group "Data Analysis and Numerical Classification" (AG DANK) for the period of 2004 – 2007.

Chairman:
Prof. Dr. Gunter Ritter, Fakultät für Mathematik und Informatik, Universität Passau, D-94030 Passau Tel.: +851/509-3110, Fax: +851/509-3112, e-mail: ritter@stoch.fmi.uni-passau.de

Vice-chairman:
Dr. Christian Hennig, Zentrum für Modellierung und Simulation, Fachbereich Mathematik, Universität Hamburg, Bundesstr. 55, D-20146 Hamburg, Tel.: +40/42838-4907, Fax: +40/42838-4924, e-mail: hennig@math.uni-hamburg.de

The German Classification Society (GfKl) will hold its 29th Annual Conference at the University of Magdeburg, March 9-11, 2005 under the title "From Data and Information Analysis to Knowledge Engineering". Further information is available from the local organizer:
Prof. Dr. Myra Spiliopoulou, Institute of Technical and Business Information Systems, Faculty of Computer Science, Otto-von-Guericke-Universitaet Magdeburg, PO Box 4120, D-39016 Magdeburg, Tel.: +391/67-18967, Fax: +391/67-18110, e-mail: myra@iti.cs.uni-magdeburg.de

Otto Opitz

New publications

In the Springer series 'Studies in Classification, Data Analysis, and Knowledge Organization'
two new volumes have recently been published:

M. Schader, W. Gaul, M. Vichi (eds.): Between Data Science and Applied Statistics. Proc. 26th Annual Conference of the GfKl, University of Mannheim, 2002. Springer, Heidelberg, 2003. 693 pp. ISBN 3-540-40354-x

This volume contains a refereed selection of 77 from the 137 papers presented during the conference. They present theoretical investigations as well as applied methods for specific practical problems. The articles are arranged in five sections:
I.   Clustering and Classification
II. Data Analysis and Statistics
III. Data Mining, Information Processing, and Automation
IV. Finance, Marketing, and Management Science
V. Biology, Archaeology, and Medicine.

H.-H. Bock, M. Chiodi, A. Mineo (eds.): Advances in Multivariate Analysis. Proc. of the Meeting of the Classification and Data Analysis Group (CLADAG)

in the Società Italiana di Statistica, University of Palermo, Italy, 2001. Springer, Heidelberg, 2004. 280pp. ISBN 3-540-20889-5
This is a refereed collection of 22 (from 54) papers presented during the Palermo meeting. Its articles are classified into five sections as follows:
I.   Classification Methods with Applications
II. Time Series Analysis and Related Methods
III. Computer Intensive Techniques and Algorithms
IV. Classification ad Data Analysis in Economics
V. Multivariate Analysis in Applied Sciences.

One further volume is in press:
D. Banks, L. Haus, F.R. McMorris, Ph. Arabie, W. Gaul (eds.): Classification, Clustering, and Data Mining Applications. Proc. of the 9th IFCS Conference, Illinois Institute of Technology, Chicago, USA, July 2004.

It will be distributed in Chicago for the conference participants

H.-H. Bock

Joint conference of the GfKl and the VOC: "Classification, the ubiquitous challenge"

The 28th Annual Conference of the GfKl, jointly organized with the VOC, took place March 9 – 11, 2004, in Dortmund (Germany). With 198 papers in 74 sessions and 2 workshops it was the biggest GfKl conference ever, even bigger than the GfKl conference 1987 in Aachen with overall 189 papers which was the 1st IFCS conference at the same time. The range of the presented topics was impressive as well. This year's conference comprised 17 invited (semi-) plenary sessions, and 57 parallel sessions divided by topic as follows:

- Classification, Clustering, and Data Analysis sessions: 12 (with 35 papers),
- Credit Risk, Capital Markets, Economics, Insurance, and Marketing sessions: 8 (26),
- Biometrics, Biostatistics, Bioinformatics, and Phylogenomics sessions: 6 (19),
- VOC organized sessions on Mixture Modelling, Optimal Scaling, Multi-way Methods and Psychometrics: 5 (18),
- Librarian sessions: 5 (10),
- Music analysis sessions: 3 (9),
- Linguistics sessions: 3 (9),
- Archaeology sessions: 3 (9),
- Sponsored sessions on Classification in Diagnostics (Roche Diagnostics), and on Pharmaceutical Research and Development (Novartis): 3 (9),
- Contest on "Social Milieus in Dortmund" sessions: 2 (6), and

- 7 other sessions with 22 papers.
Moreover, two workshops on "Multimedia in Education" and "Statistics in Media" took place with 9 papers altogether.

280 paying participants joined the conference, 184 from Germany, 21 from the Netherlands, 19 from Poland, 13 from Austria, 9 from Switzerland, 8 from Belgium, 6 from the U.S.A., and 20 from 11 other countries. Meanwhile around 90 papers were submitted for the conference proceedings to be published by Springer next year.

Claus Weihs

Other news

The traditional Fall Meeting 2004 of the Working Group ``Data Analysis and Numerical Classification'' (AG-DANK) will take place at the University of Marburg; the meeting is invited and organized by Professor Alfred Ultsch. The focus this year will be on "Analysis of time series, in particular musical time series". Further information will be available at: http://stoch.fmi.uni-passau.de//agdank/Marburg2004/

Gunter Ritter
ritter@fmi.uni-passau.de

NEWS FROM THE SFC

Elections
Rosanna VERDE (Seconda Universita' Degli Studi Di Napoli) has been elected as « Vice-President » of the SFC. Djamel Abdelkader ZIGHED (Laboratoire ERIC, Université Lumière Lyon 2) has been elected as "Vice-President" of the SFC.

SFC' 2004 Conference, September 8-10, Bordeaux, France
The *"11èmes rencontres de la Société Francophone de Classification"*, yearly meeting of the SFC, will take place in Bordeaux, France, from September 8 to September 10, 2004. This conference is organized by the "*Laboratoire de Mathématiques Appliquées de Bordeaux*" (MAB) (Bordeaux 1 and Bordeaux 2 Universities) and it will be located at the Mathematical Institute of Bordeaux on the campus of Talence. Prof. Michel Langlais is the chairman of the scientific program committee.
Together with the usual topics of classification and clustering, the main themes of the conference are: Pattern Recognition, Environment, Data Mining, Genomic, Scoring, Symbolic Data Analysis, Taxonomic Cognitive Models.

For further details, members of the scientific committee, organizers, invited speakers, accommodation facilities, deadlines, please visit our website: http://www.math.u-bordeaux.fr/SFC04
e-mail : SFC04@math.u-bordeaux.fr

12th Annual meeting of the SFC
SFC'2005 meeting will take place in Montreal from May 30 to June 1, 2005

André HARDY, SFC Secretary

NEWS FROM THE VOC

The VOC co-organised a successful joint meeting with the GFKL from March 9 to 11 in Dortmund, Germany. Willem Heiser, editor of the Journal of Classification and one of the founders of the VOC, gave the opening address on correspondence analysis. The current IFCS president and past chairman of the VOC Henk Kiers presented an invited paper on bootstrap confidence intervals for three-way component models. Furthermore, a number of VOC members organised four interesting sessions on optimal scaling (A. van der Kooij and E. Dusseldorp), mixture modelling (J. Vermunt), psychometric methods (P. de Boeck) and three-way methods (H. Kiers).

The 15th anniversary of the VOC will be celebrated in Driebergen, The Netherlands on November 11 and 12, 2004. The theme of the meeting is "All in Good Place", that is, on spatial data and spatial representation of data. The program includes distinguished speakers as T. Bijmolt, A. Buja, H. Elffers, J. de Gruyter, W. Heiser, S. Heisterkamp, H. Kiers, E. Le Saffre, K. van Deun, M. van de Velden and R. Wehrens. All interested are kindly invited to attend this meeting. The up-to-date information on the activities of the VOC is available via www.voc.ac.

Marieke Timmerman

MESSAGE FROM JOURNAL OF CLASSIFICATION

In the article section of present issue, the first three papers all concern classical clustering methodology. Gascuel and McKenzie study the several well-known hierarchical clustering algorithms in terms of their safety radius, a measure that describes worst case behavior. They demonstrate that the usual methods for hierarchical trees have safety radius twice as large as the radius of usual methods for additive trees. Remarkably, they also show that least squares methods are quite vulnerable to data perturbations

when the number of objects becomes large. In the context of the HICLASS model, which describes hierarchical clusters in binary data, Ceulemans, Van Mechelen, and Kuppens give some nice extensions of earlier work on the three-way case. Their paper offers a taxonomy of a wide range of model constraints, shows how to deal with them algorithmically, and discusses illuminating examples from psychiatric diagnosis and personality research. In the third paper, Gregorius discusses a principle of hierarchical clustering based on classical ideas of internal cohesion and external isolation. He shows that the chosen criterion of isolation—which characterizes a model, not an algorithm—always and unambiguously generates a set of nested clusters.

Then there are two dimension reduction papers. Using a projection pursuit framework, Montanari proposes a new criterion for linear discriminant analysis, based on Gini's transvariation probability as a projection index. She shows that this discriminant function is robust when dealing with non-normal and heteroscedastic distributions, and reduces to Fisher's linear discriminant function when the subgroup distributions are regular. Next, Van de Velden revisits Guttman's method of quantifying paired comparison data and shows that there is an alternative type of analysis, very similar yet basically different from the usual one, which optimally differentiates between subjects (or other sources from which repeated paired comparisons were obtained).

We conclude with two statistical papers. Song and Lee consider mixtures of structural equation models, and develop local influence methodology to assess whether results are sensitive to minor perturbations of model inputs, such as data deletion and changes in model specifications. Finally, Dodge and Lemburger give general asymptotic properties of estimators for several centers of multivariate density functions. Apart from the usual multivariate mean, the centers include several types of multivariate medians, which are shown to be preferable to the mean in specified circumstances.

Willem J. Heiser

CALL FOR PAPERS - EGC 2005

The 5th Conference on Extraction and Knowledge Management January 19-21, 2005
To be held at Université Paris 5, René Descartes, PARIS 6, FRANCE
mail : egc2005@math-info.univ-paris5.fr

Honorary chair : Gregory Piatetsky-Shapiro - President of KDNuggets, Brookline, MA, USA
Co-chairs : Suzanne Pinson – Université Paris 9-Dauphine and Nicole Vincent – Université Paris 5-René Descartes.
Steering Committee:
H. Briand (LINA Nantes), D. Boulanger (Lyon3), R. Gras (LINA Nantes), F. Guillet (LINA Nantes), M.S. Hacid (LIRIS Lyon), D. Hérin (LIRMM Montpellier), Y.Kodratoff (LRI Paris), L. Lebart (ENST Paris), J. Phillipé (LINA Nantes), D. Zighed (ERIC Lyon 2).

Important dates :
Submission deadline : October 1st, 2004; Notification of acceptance : November 14th, 2004; Camera Ready version : November 24th, 2004

Conference aims
The Conference aims to bring together academic and industrial researchers in various fields related to those topics such as Knowledge Bases, Learning, Knowledge Discovery, Knowledge Representation, Data, Text and Web Mining, Data Warehouse. The conference will address recent advances in theory, methodologies and applications in these different fields.
Specific topics include, but are not limited to: Knowledge discovery, Data and knowledge modelling, Data and knowledge representation (data warehouse), Knowledge bases, Ontologies (acquisition, update, interoperability), Data and model visualisation, data and image mining, Agent-mediated Knowledge Management, Business Intelligence, Data and Knowledge Acquisition, pre-processing and integration, Methods and algorithms for knowledge extraction, Very large and complex data processing (texts, images, videos, audio, spatial and temporal data, semi-structured data, Web mining, Web services). Application domains : industry, administration, transportation, tourism, defence, biology, telecommunications, etc.

Other features
A software demo session during the conference; Workshops and tutorials the day before the conference; An award given to the best paper during a special session: EGC best paper award; 15 scholarships attributed to PhD students with an accepted paper.

Submission procedure
Theoretical and applied contributions are invited in the form of long papers (max. 12 pages) for original works or short papers (max. 6 pages) for work in progress following the formatting style given on the Web site of the conference. The conference language is French but papers can be submitted in English.

CALL FOR PAPERS: ECML/PKDD04

Workshop on Symbolic and Spatial Data Analysis : Mining Complex Data Structures
SEPTEMBER 20, 2004, PISA – Italia

Workshop web site:
<http://www.info.fundp.ac.be/ssda-pkdd04/>http://www.info.fundp.ac.be/ssda-pkdd04/
Conference web site: <http://ecmlpkdd.isti.cnr.it/>http://ecmlpkdd.isti.cnr.it/

IMPORTANT DATES
Submission deadline: June 14, 2004, Notification of acceptance: July 5, 2004, Camera-ready copies due: July 12, 2004.

WORKSHOP CHAIRS
Monique Noirhomme-Fraiture (Institut d'Informatique, Namur - Belgium); Paula Brito (Faculdade de Economia, University of Porto)

OBJECTIVE
The main goal of this workshop is to bring together researchers from different communities such as machine learning, data analysis, symbolic data analysis and data mining to promote discussion and the development of new ideas and methods to deal with complex structured data.

WORKSHOP STRUCTURE AND ATTENDANCE
The workshop is intended to be a highly communicative meeting place for researchers working on similar topics, but coming from different communities. In order to achieve these goals, the workshop will consist of invited talks, short presentations and discussions. Depending on the number of attendees, participants may be asked to present themselves, and state their research interests. Depending on the submitted papers, sessions will be organized by topics. At the end of each session, a discussion of the presented papers will take place.

TOPICS
The workshop will aim at a balance between theoretical issues and descriptions of case studies. Topics of interest include but are not limited to:
Classification with Structured Data, Complex Data Analysis, Data Summarizing and Data Mining, Mining Aggregated Data Spatial Data Analysis, Symbolic Data Analysis, Visualization of Complex Data

SUBJECT
SYMBOLIC DATA ANALYSIS
There is an increasing need to extend standard exploratory, statistical and graphical data analysis methods to the case of more complex data, that go beyond the classical framework. This is the case of data concerning more or less homogeneous classes or groups of individuals (second-order objects or macro-data), instead of single individuals (first-order objects or micro data). The extension of classical data analysis techniques to the analysis of second-order objects is one of the main goals of a novel research field names "symbolic data analysis". Symbolic data extend the classical tabular model, allowing multiple, possibly weighted, values for each descriptive attribute which allow representing variability and/or uncertainty present in the data. Symbolic Data Analysis methods include unvaried descriptive methods, clustering, decision-tree, discrimination, regression and factorial analysis techniques, which allow analyzing symbolic data tables. Symbolic data occur in many situations, for instance in summarizing huge sets of data or in describing the underlying concepts (a town, a socio-demographic group, a scenario of accidents) of a database. It also finds an important application field in official statistics; since by law, NSI's are prohibited from releasing individual responses to any other government agency or to any individual or business, data are aggregated for reasons of privacy before being distributed to external agencies and institutes. Symbolic Data Analysis provides useful tools to analyze such aggregated data. Symbolic Data Analysis underwent great improvement with the European projects "Symbolic Data Analysis Systems (SODAS)" and "Analysis System for Symbolic Official Data (ASSO)"; as the result of these projects a software package SODAS has been developed.

SPATIAL DATA ANALYSIS
With the exponentially growing use of geographic information systems (GIS) to store, manipulate and visualize geo-coded information, it is increasingly important to understand the particular nature of geographic data and the specialized techniques required for its analysis. There is increasing interest in studying how techniques for the analysis of spatial data can be effectively applied in a GIS environment, such as the study of spatial patterns and spatial autocorrelation, detection of clusters, outliers and any other relationships that pertain to the absolute and relative location of observations. Common applications of spatial data analysis techniques in the social sciences range from the discovery of crime clusters, hot spots and the detection of disease clusters, to spatial autocorrelations of demographic variables and regression models for real estate analysis. Other applications concern public health services searching for explanations of disease clusters, environmental agencies assessing the impact of changing land use patterns on climate change, geo-marketing companies doing customer segmentation based on spatial location, etc. For supporting this type of analysis, most contemporary GIS have only very basic spatial analysis functionalities; many are confined to analysis that involves descriptive statistical displays, such as histograms or pie charts. Data mining, which is the partially automated search for hidden patterns in large databases, offers great potential benefits for the applied GIS based decision-making that takes place in public and private sector organizations.

SUBMISSIONS
See also WEB SITE
http://www.info.fundp.ac.be/ssda-pkdd04/
Submission Procedure
Authors are invited to submit original research contributions or experience reports in English. Submitted papers must be unpublished and substantially different from papers under review. Papers that have been or will be presented at small workshops/symposia whose proceedings are available only to the attendees may be submitted. Papers should be double-spaced and no longer than 5000 words (about 12 single-spaced pages). Papers should be sent electronically (postscript or pdf) not later than June 14, 2004.
Papers will be selected on the basis of review of full paper contributions. Notification of acceptance will be given by July 5, 2004. Final camera-ready copies of accepted papers will be due by July 12, 2004. The proceedings will be printed by the ECML/PKDD organizers and distributed at the workshop. A web-publication of the proceedings is expected after the conference. It is intended to publish a selection of papers as a special number of an international journal.

Style Guide
There is a joint paper style for the proceedings of all ECML/PKDD workshops. Submitted papers should be formatted according to the Springer-Verlag Lecture Notes in Artificial Intelligence guidelines. Authors' instructions and style files can be downloaded from http://www.springer.de/comp/lncs/authors.html.

PROGRAM COMMITTEE
Annalisa Appice, Patrice Bertrand, Lynne Billard, Gilles Bisson, Hans-Hermann Bock, Carmen Bravo, Marie Chavent, Francisco De Carvalho, Edwin Diday, K. Chidananda Gowda, Georges Hebrail,

Manabu Ichino, Yves Lechevallier, Michèle Sebag, Petko Valtchev, Rosanna Verde.

CONTACT
Paula Brito, Faculdade de Economia, University of Porto, Rua Dr. Roberto Frias, 4200-464 Porto, PORTUGAL ; Tel : (+351) 225571233; Fax : (+351) 225505050, e-mail: mpbrito@fep.up.pt
Monique Noirhomme-Fraiture, Institut d'Informatique, Facultés Universitaires Notre Dame de la Paix, Rue Grandgagnage, 21, B-5000 Namur , BELGIUM ; Tel : (+32) 81724979; Fax : (+32) 81724967 ; e-mail: monique.noirhomme@info.fundp.ac.be

Contributions for the coming issue of the IFCS newsletter can be sent to:

krzysztof.jajuga@ae.wroc.pl

Text files are by preference in ascii or word, with pc format. For graphical materials, by preference GIF is used.