07181 Abstracts Collection
Parallel Universes and Local Patterns
— Dagstuhl Seminar —

Michael R. Berthold[1], Katharina Morik[2] and Arno Siebes[3]

[1] Univ. Konstanz, DE
berthold@inf.uni-konstanz.de
[2] Univ. Dortmund
morik@ls8.informatik.uni-dortmund.de
[3] Utrecht Univ., NL
siebes@cs.uu.nl

Abstract. From 1 May 2007 to 4 May 2007 the Dagstuhl Seminar 07181 "Parallel Universes and Local Patterns" was held in the International Conference and Research Center (IBFI), Schloss Dagstuhl. During the seminar, several participants presented their current research, and ongoing work and open problems were discussed. Abstracts of the presentations given during the seminar as well as abstracts of seminar results and ideas are put together in this paper. The first section describes the seminar topics and goals in general. Links to extended abstracts or full papers are provided, if available.

Keywords. Local Patterns, Global Models, Parallel Universes, Descriptor Spaces

07181 Introduction – Parallel Universes and Local Patterns

Learning in parallel universes and the mining for local patterns are both relatively new fields of research. Local pattern detection addresses the problem of identifying (small) deviations from an overall distribution of some underlying data in some feature space. Learning in parallel universes on the other hand, deals with the analysis of objects, which are given in different feature spaces, i.e. parallel universes; and the aim is on finding groups of objects, which show "interesting" behavior in some of these universes. So, while local patterns describe interesting properties of a subset of the overall space or set of objects, learning in parallel universes also aims at finding interesting patterns across different feature spaces or object descriptions. Dagstuhl Seminar 07181 on Parallel Universes and Local Patterns held in May 2007 brought together researchers with different backgrounds to discuss latest advances in both fields and to draw connections between the two.

Keywords: Local Patterns, Global Models, Parallel Universes, Descriptor Spaces

Dagstuhl Seminar Proceedings 07181
Parallel Universes and Local Patterns
http://drops.dagstuhl.de/opus/volltexte/2007/1266

Joint work of: Berthold, Michael R.; Morik, Katharina; Siebes, Arno

Full Paper: http://drops.dagstuhl.de/opus/volltexte/2007/1265

Local Patterns in Plastic Card Fraud Detection

Niall Adams (Imperial College London, GB)

We discuss local patterns in the context of plastic card transaction fraud detection.

Keywords: Local patterns, fraud detection

Note on parallel universes

Niall Adams (Imperial College London, GB)

The parallel universes idea is an attempt to integrate several aspects of learning which share some common aspects. This is an interesting idea: if successful, insights could cross-fertilise, leading to advances in each area. The "multi-view" perspective seems to us to have particular potential.

Keywords: Local patterns, fraud detection

Joint work of: Adams, Niall; Hand, David J.

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2007/1256

Subspace outlier mining in large multimedia databases

Ira Assent (RWTH Aachen, D)

Increasingly large multimedia databases in life sciences, e-commerce, or monitoring applications cannot be browsed manually, but require automatic knowledge discovery in databases (KDD) techniques to detect novel and interesting patterns. One of the major tasks in KDD, clustering, aims at grouping similar objects into clusters, separating dissimilar objects. Density-based clustering has been shown to detect arbitrarily shaped clusters even in noisy data bases.

In high-dimensional data bases, meaningful clusters can no longer be detected due to the "curse of dimensionality". Consequently, subspace clustering searches for clusters hidden in any subset of the set of dimensions. As the number of subspaces is exponential in the number of dimensions, traditional approaches use fixed pruning thresholds. This results in dimensionality bias, i.e. with growing dimensionality, more clusters are missed.

Clustering information is very useful for applications like fraud detection where outliers, i.e. objects which differ from all clusters, are searched. In subspace

clustering, an object may be an outlier with respect to some groups, but not with respect to others, leading to possibly conflicting information.

We propose a density-based unbiased subspace clustering model for outlier detection. We define outliers with respect to all maximal and non-redundant subspace clusters, taking their distance (deviation in attribute values), relevance (number of attributes covered) and support (number of objects covered) into account.

We demonstrate the quality of our subspace clustering results in experiments on real world and synthetic databases and discuss our outlier model.

Keywords: Data mining, outlier detection, subspace clustering, density-based clustering

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2007/1257

Interactive, supervised learning in Parallel Universes

Michael R. Berthold (Universität Konstanz, D)

Classical data analysis methods typically assume that all objects of a data set are described in a single feature space. This feature space is assumed to comprise all necessary information to classify an object. However, in many real-world applications there are numerous ways to describe complex objects. An example are musical songs, i.e. audio streams, which can be represented based on dynamics, melody, and key or - as a different representation - based on rhythm and harmony. A third representation may be more descriptive, such as interpreter, position in music charts, length, and so on. Further examples of complex objects are images, 3D objects or molecules in drug discovery. With regard to learning, such as clustering or building classification models, it is often unclear, which of the available descriptors are optimal for any given task. Clustering in Parallel Universes is a new research field that deals with such multiply described data sets. It aims at identifying interesting patterns in data, e.g. groups of objects that cluster well in one (or few) universe(s).

We describe a supervised, interactive way to generate a set of clusters for a given data set in Parallel Universes. The clustering is done by constructing local histograms for each object of a (set of) target class(es), which can then be used to visualize, select, and fine-tune potential cluster candidates across different universes.

The accompanying algorithm can also generate clusters automatically by assigning a quality value to each Neighborgram and greedily adding the best Neighborgram, no matter from which universe it stems, to the global prediction model. The procedure enables an automatic or semi-automatic clustering process where the user only occasionally interacts with the algorithm.

Keywords: Parallel universes

Learning in Parallel Universes

Michael R. Berthold (Universität Konstanz, D)

This abstract summarizes a brief, preliminary formalization of learning in parallel universes. It also attempts to highlight a few neighboring learning paradigms to illustrate how parallel learning fits into the greater picture.

Keywords: Parallel universes

Joint work of: Berthold, Michael R.; Wiswedel, Bernd

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2007/1258

Trajectories in Parallel Universes

Francesco Bonchi (ISTI-CNR - Pisa, I)

In this talk we will present two recent results on spatio-temporal data.

The first work tackles the problem of anonymity when publishing a database of trajectories, by means of a clustering-based approach. The second work introduces a brand new kind of trajectory local pattern.

We will also discuss possible relations of these two works with the local patterns and parallel universes frameworks.

Don't be afraid of so-many local patterns

Jean-François Boulicaut (INSA - Lyon, F)

We have been working on local pattern discovery for a while (constraint-based mining of set patterns from 0/1 data, constraint-based mining of sequential pattern from sequence databases). More recently, we have investigated a local to global framework that leads to a new approach for bi-clustering 0/1 data or, more generally, knowledge discovery based on bi-clustering. In this short talk, we will discuss our view on constraint-based data mining (aka inductive querying) by considering a couple of simple but yet useful KDD processes. This will help to discuss our research agenda for taking the most from (possibly huge) collections of local patterns.

Keywords: Clustering, bi-clustering, local pattern mining, gene expression data analysis

Discovering Knowledge from Local Patterns with Global Constraints

Bruno Crémilleux (Université de Caen, F)

It is well known that local patterns are at the core of a lot of knowledge which may be discovered from data.

Nevertheless, use of local patterns is limited by their huge number and computational costs. Several approaches (e.g., condensed representations, pattern set discovery) aim at grouping or synthesizing local patterns to provide a global view of the data. A global pattern is a pattern which is a set or a synthesis of local patterns coming from the data. In this paper, we propose the idea of global constraints to write queries addressing global patterns. A key point is the ability to bias the designing of global patterns according to the expectation of the user. For instance, a global pattern can be oriented towards the search of exceptions or a clustering. It requires to write queries taking into account such biases. Open issues are to design a generic framework to express powerful global constraints and solvers to mine them. We think that global constraints are a promising way to discover relevant global patterns.

Keywords: Local patterns, constraint-based paradigm, global constraints

Joint work of: Crémilleux, Bruno; Soulet, Arnaud

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2007/1259

Mining Optimal Decision Trees from Itemset Lattices

Elisa Fromont (Katholieke Universiteit Leuven, B)

We present an exact algorithm for finding a decision tree that optimizes a ranking function under size, depth, accuracy and leaf constraints. Because the discovery of optimal trees has high theoretical complexity, until now no efforts have been made to compute such trees for real-world datasets. An exact algorithm is of both scientific and practical interest. From a scientific point of view, it can be used as a gold standard to evaluate the performance of heuristic decision tree learners and to gain new insight in these traditional learners. From the application point of view, it can be used to discover trees that cannot be found by heuristic decision tree learners. The key idea behind our algorithm is the relation between constraints on decision trees and constraints on itemsets. We propose to exploit lattices of itemsets, from which we can extract optimal decision trees in linear time. We give several strategies to efficiently build these lattices and show that the test set accuracies of C4.5 compete with the test set accuracies of optimal trees, which confirms the common assumption that heuristic decision tree learners usually identify trees that generalize very well.

Keywords: Decision tree learning, Frequent itemset mining, Constraint based mining

Joint work of: Fromont, Elisa; Nijssen, Sigfried

Parallel universes to improve the diagnosis of cardiac arrhythmias

Elisa Fromont (Katholieke Universiteit Leuven, B)

We are interested in using parallel universes to learn interpretable models that can be subsequently used to automatically diagnose cardiac arrythmias. In our study, parallel universes are heterogeneous sources such as electrocardiograms, blood pressure measurements, phonocardiograms etc. that give relevant information about the cardiac state of a patient. To learn interpretable rules, we use an inductive logic programming (ILP) method on a symbolic version of our data. Aggregating the symbolic data coming from all the sources before learning, increases both the number of possible relations that can be learned and the richness of the language. We propose a two-step strategy to deal with these dimensionality problems when using ILP. First, rules are learned independently in each universe. Second, the learned rules are used to bias a new learning process from the aggregated data. The results show that this method is much more efficient than learning directly from the aggregated data. Furthermore the good accuracy results confirm the benefits of using multiple sources when trying to improve the diagnosis of cardiac arrythmias.

Keywords: Parallel universes, inductive logic programming, medical application, declarative bias

Joint work of: Fromont, Elisa; Quiniou, Rene; Cordier, Marie-Odile

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2007/1260

Mining Frequent Itemsets in Streaming Data

Bart Goethals (University of Antwerp, B)

We introduce a new frequency measure for itemsets in a stream, present an algorithm to find all frequent itemsets and study some of its interesting features using well known results from number theory.

Reliably Capture Local Clusters in Noisy Domains From Parallel Universes

Frank Höppner (FH Wolfenbüttel, D)

When seeking for small local patterns it is very intricate to distinguish between incidental agglomeration of noisy points and true local patterns.

We propose a new approach that addresses this problem by exploiting temporal information which is contained in most business data sets. The algorithm enables the detection of local patterns in noisy data sets more reliable compared to the case when the temporal information is ignored. This is achieved by making use of the fact that noise does not reproduce its incidental structure but even small patterns do. In particular, we developed a method to track clusters over time based on an optimal match of data partitions between time periods.

Keywords: Local pattern, time, parallel universe

Joint work of: Höppner, Frank; Böttcher, Mirko

Extended Abstract: http://drops.dagstuhl.de/opus/volltexte/2007/1261

Finding Orderings and Reverse Orderings in Noisy Data

Frank Klawonn (FH Wolfenbüttel, D)

Various high throughput technologies (e.g. DNA microarrays, gas and mass spectrometry) enable biologists to access important classes of cellular bio-molecules in order to gain insight into the corresponding biological processes. Unfortunately, measurements from these high throughput technologies tend to be affected by a high influence of noise, making it more difficult to discover local patterns and clusters in the data. In our case, the patterns correspond to orders. Biological "items" (like metabolites or genes) showing the same behaviour or the opposite behaviour (in terms of order) w.r.t. various biological conditions or over a period of time are of interested. We try to find these interesting order patterns taking the noisy measurements into account based on statistical models for the estimation of the noise.

Building Global Models from Local Patterns

Arno J. Knobbe (Utrecht University, NL)

In this talk I will outline a model for Knowledge Discovery that we have been using. The model is based on building larger structures that integrate locally discovered models. It is centred around two main phases: Pattern Discovery and Pattern Combination. The first phase is responsible for producing a, typically large, collection of local patterns in a relatively untargeted and exploratory manner (typically using Subgroup Discovery). In the second phase then, patterns are selected and combined to form global models. In order for this process to work effectively and obtain a good division of work, information needs to be exchanged between the two phases.

Keywords: Local Patterns, Subgroup Discovery, Pattern Teams

Visualization in Complex Domains: Graphical Model Introspection

Rudolf Kruse and Matthias Steinbrecher (Universität Magdeburg, D)

The production pipeline of present day's automobile manufacturers consists of a highly heterogeneous and intricate assembly workflow that is driven by a considerable degree of interdependencies between the participating instances as there are suppliers, manufacturing engineers, marketing analysts and development researchers. Therefore, it is of paramount importance to enable logistics experts to quickly respond to potential on-time delivery failures, ordering peaks or other disturbances that may interfere with the ideal assembly process.

These requirements call for treatment methods that exploit the dependence structures embedded inside the application domain. Furthermore, these methods need to be equipped with dedicated updating, revision and refinement techniques in order to cope with the above-mentioned possible irregularities.

Since every production and planning stage involves highly specialized domain experts, it is necessary to offer intuitive system interfaces that are less prone to inter-domain misunderstandings.

The presentation refers to industrial details and an alternative interpretation of Bayesian networks — as one type of graphical model — that results in an easily assessable visualization method that helps to find interesting patterns inside huge data volumes.

Keywords: Graphical Models, Visualization

References

1. Gebhardt, J., Detmer, H., Madsen, A.L.: Predicting Parts Demand in the Automotive Industry - An Application of Probabilistic Graphical Models. In: Proc. Int. Joint Conf. on Uncertainty in Artificial Intelligence (UAI'03, Acapulco, Mexico), Bayesian Modelling Applications Workshop. (2003)

2. Steinbrecher, M., Kruse, R.: An Alternative Interpretation of Probabilistic Potentials for Exploratory Data Analysis. In: Proc. 1st Joint Conf. German Statistical Society. (2007)

3. Borgelt, C., Kruse, R.: Graphical Models - Methods for Data Analysis and Mining. John Wiley & Sons, United Kingdom (2002)

Finding all Local Models in Parallel: Multi-Objective SVM

Ingo Mierswa (Universität Dortmund, D)

Recently, evolutionary computation has been successfully integrated into statistical learning methods.

A Support Vector Machine (SVM) using evolution strategies for its optimization problem frequently deliver better results with respect to the optimization criterion and the prediction accuracy. Moreover, evolutionary computation allows for the efficient large margin optimization of a huge family of new kernel functions, namely non-positive semidefinite kernels as the Epanechnikov kernel. For these kernel functions, evolutionary SVM even outperform other learning methods like the Relevance Vector Machine. In this paper, we will discuss another major advantage of evolutionary SVM compared to traditional SVM solutions: we can explicitly optimize the inherent trade-off between training error and model complexity by embedding multi-objective optimization into the evolutionary SVM. Hence, it is no longer necessary to tune the SVM parameter C which weighs both conflicting criteria. The result is the complete Pareto front starting with the global model alone, adding all local models and ending in the overfitting case.

The user can actually see the point where overfitting occurs and can easily select a solution from the Pareto front best suiting his or her needs. The major result of this work is that this complete knowledge about the different models can be derived in parallel, i.e. in one single run of the SVM.

Keywords: Support vector machines, machine learning, kernel methods, evolution strategies, local models

Multi-Aspect Tagging for Collaborative Structuring

Katharina Morik (Universität Dortmund, D)

Local tag structures have become frequent though Web 2.0:

Users "tag" their data without specifying the underlying semantics.

A collection of media items is tagged multiply using different aspects, e.g., topic, genre, occasion, mood. Given the large number of local, individual structures, users could benefit from the tagging work of others ("folksonomies"). In contrast to distributed clustering, no global structure is wanted. Each user wants to keep the tags already annotated, wants to keep the diverse aspects under which the items were organized, and only wishes to enhance the own structure by those of others.

A clustering algorithm which structures items has to take into account the local, multi-aspect nature of the task structures.

The LACE algorithm (Wurst et al. 2006) is such a clustering algorithm.

Keywords: Ensemble Clustering, automatic tagging, localized clustering

Joint work of: Morik, Katharina; Wurst, Michael

Full Paper:
http://www-ai.cs.uni-dortmund.de/auto?self=%24Publication%5feoqyevw7

See also: Wurst, M. and Morik, K. and Mierswa, I. Localized Alternative Cluster Ensembles for Collaborative Structuring. In Proc. of the European Conference on Machine Learning (ECML), 2006.

Parallel universes and local patterns for learning regulation graphs : a case study

Céline Rouveirol (Université Paris Nord, F)

Parallel universes and local patterns for learning regulation graphs : a case study

Gene regulation in eukaryotes involves many complex mechanisms, most of which are not well understood. With the advent of high-throughput microarray technologies, the expression levels of thousands of genes can be measured simultaneously during various biological processes and for collections of related samples. Considerable effort has been devoted to the analysis of these data sets for the reconstruction of regulatory networks. A family of approaches based on mathematical models of the regulation process has been developed (e.g. Boolean, Bayesian, piecewise-linear, probabilistic Boolean, ...). Attempts to learn such models from expression data are hindered by the large number of potential solutions, and the unrealistically large amount of data required to identify the best solution. In cases of complex formalism for the modelling of regulation, in particular, it has only been possible to reconstruct subnetworks with a few variables.

Considerable effort is currently being dedicated to the charting of large-scale gene regulatory networks, relating the expression of a target gene to that of the genes encoding its regulators.

Recent integrative studies have aimed to derive complete yeast gene networks from additional information (e.g. protein-DNA binding from ChIP-chip experiments or computational analysis of transcription factor binding sites), with the computational advantage of restricting the number of possible regulators for a given target gene. However, these approaches are difficult to adapt to other organisms, for which the computational detection of cis-elements is more difficult, and the experimental detection of binding events is currently limited (e.g. Homo sapiens). In contrast, expression data sets are being collected at an exponential rate, and methods based solely on the use of gene expression for network reconstruction are required.

(Peer et al., 2002) have designed the Minreg system, a constrained Bayesian network for the reconstruction of large-scale regulatory networks from expression data. The maximal in-degree (i.e., the number of regulators) of target genes and the total number of regulators in the model are limited, so the model focuses on only a small set of global active regulators Ra. The authors made use of these constraints to devise an approximation algorithm for searching for high scoring networks among expression data. The system successfully and robustly identifies the key active regulators, but cannot learn the full detailed network, and may

miss interesting regulation relationships: given a current set of active regulators Ra, the greedy search of Minreg will ignore combinations of co-regulators $R_a \cap \{r_1, r_2\}$ if the marginal score values of $Ra \cap \{r_1\}$ and $R_a \cap \{r_2\}$ are both low, although $R_a \cap \{r_1, r_2\}$ may be significant. In such a case, r_1 and r_2 are said to co-operate — i.e. they act collectively to influence their target genes. Previous computational approaches, due to complexity reasons, have therefore only partly investigated the role of regulator co-operativity. However, such mechanisms have been identified in many organisms (e.g. Saccharomyces cerevisiae, Homo sapiens).

We have proposed an original, scalable technique called LICORN (for Learn-Ing CO-operative Regulation Networks) for deriving co-operative regulations, in which many co-regulators act together to activate or repress a target gene. Many forms of combinatorial logical control may theoretically occur in Boolean or Bayesian models, but we focus here on co-operative regulation patterns that i) follow the biologically justified activator-repressor model ii) operate on ternary expression level representation iii) allow for efficient large-scale network computation. LICORN uses an original heuristic approach to accelerate the search for an appropriate structure for the regulation network. It first extracts a global, condensed representation of frequent co-regulator sets using constrained item-set mining techniques. From this representation, a limited subset of candidate co-regulator sets is then efficiently associated with each gene. As this candidate subset is modest in size, exhaustive searches for the best gene regulatory network can be performed.

Given a set of target genes G, a set of regulators R, discretised expression matrices MG and MR for genes and regulators over the set samples and an evaluation score h, associating a real number with a candidate Gene Regulation Network (GRN), our goal is to find, for each target gene g, the set of regulators that best explains the level of expression of g.

The first step of LICORN generates a set of candidate co-regulator sets for all genes of G, such that a candidate co-regulator set is a set of regulators frequently co-expressed in the data. During the second step, for each target gene of G, LICORN efficiently computes a limited set of candidate GRNs and then exhaustively searches for the best one in this set — the activator and inhibitor sets best explaining the target gene status in the sample set. The last step of LICORN is a permutation-based method for the selection of statistically significant GRNs from the inferred GRNs for all target genes.

Once these significant GRNs have been inferred (providing many "local" and maybe overfitting views about how the regulation of each gene takes place), the problem is then to build one or more global view of the regulation network, to infer pathways or modules, i.e., focusing on genes and their regulators that take part in a common function. we will briefly give in our talk some hints for building this global view from these many local views.

Keywords: Bioinformatics, transcriptome data

Joint work of: Rouveirol, Céline; Elati, Mohamed

Re-ranking Subgroups for Fraud Detection

Stefan Rüping (Fraunhofer IAIS - St. Augustin, D)

This talk discusses the application of local patterns to fraud detection by identifying and re-ranking interesting subgroups.

Sequential Supervised Local Pattern Detection

Martin Scholz (Universität Dortmund, D)

In many domains it is easy to learn a rough model covering the global aspects, but the truely interesting aspects are the smaller local patterns. This talk describes a generic approach for supervised settings that allows to discover local patterns in the presence of global patterns under different measures of interestingness. It will also be pointed out how to combine such patterns for predictive purposes and how to identify patterns that are "local" in the time dimension, only. Some practical local pattern mining tasks illustrate the relevance and growing interest of these methods.

Patterns and Parallel Universes

Arno Siebes (Utrecht University, NL)

The key idea of parallel universes is that different similarity measures give a different view on the data. Molecules that are similar as far as valence is concerned may be rather different if we focus on weight.

Patterns induce a simple similarity measure, two tuples either satisfy the same pattern or they don't. So, different patterns give different views on the data. The frequent pattern explosion, however, implies an explosion of the number of, crude, parallel universe. In this talk I will discuss how our recent work on MDL for pattern mining solves this problem.

Discovery of Structured Rules

Einoshin Suzuki (Kyushu University, J)

In this talk, I will try to convince you that discovery of structured rules is a promising framework for parallel universe for rule discovery. In rule discovery, it is well-known that truly interesting rules tend to exhibit low supports but rules with low supports are huge in number and most of them are uninteresting. Methods which aim at circumventing this problem can be classified into objective approach and subjective approach, the former typically employs an

interestingness measure [GPS 91, Smyth 92, Hajek 66] while the latter employs rule templates [Klemettinen 94] or uses user-given rules [Silberstchatz 96, Padmanabhan 98, B. Liu 97, 99a]. Our structured rule discovery approach, which dates back to [Suzuki 96], belongs to the objective approach and tackles the problem by setting the target local pattern to a rule pair which consists of a general rule and its exception rule. It can be viewed a parallel universe for rule discovery because the interestingness of a rule $Y \rightarrow x$ is gauged from its subrules $Y \rightarrow x$, where each of Y and Y is a conjunction of (attribute = value)s, Y is a sub-expression of Y, each of x and x is an (attribute = value), and x and x have the same attribute. I will explain several theoretical and practical results for structured rule discovery [Suzuki 97, 00] and I will compare it with related methods including subgroup discovery [Hoschka 91, Klosgen 96], direction setting rule [B. Liu 99b], and Yugami's interesting rule [Yugami 00].

References

GPS 91 Gregory Piatetsky-Shapiro: Discovery, Analysis, and Presentation of Strong Rules Knowledge Discovery in Databases 1991: 229-248

Smyth 92 Padhraic Smyth, Rodney M. Goodman: An Information Theoretic Approach to Rule Induction from Databases. IEEE Trans. Knowl. Data Eng. 4(4): 301-316 (1992)

Hajek 66 Hajek P. - Havel I. - Chytil M.: The GUHA method of automatic hypotheses determination, Computing 1(1966) 293Ũ308.

Klemettinen 94 Mika Klemettinen, Heikki Mannila, Pirjo Ronkainen, Hannu Toivonen, A. Inkeri Verkamo: Finding Interesting Rules from Large Sets of Discovered Association Rules. CIKM 1994: 401-407

Silberschatz 96 Abraham Silberschatz, Alexander Tuzhilin: What Makes Patterns Interesting in Knowledge Discovery Systems. IEEE Trans. Knowl. Data Eng. 8(6): 970-974 (1996)

Padmanabhan 98 Balaji Padmanabhan, Alexander Tuzhilin: A Belief-Driven Method for Discovering Unexpected Patterns. KDD 1998: 94-100

B. Liu 97 Bing Liu, Wynne Hsu, Shu Chen: Using General Impressions to Analyze Discovered Classification Rules. KDD 1997: 31-36

B. Liu 99a Bing Liu, Wynne Hsu, Lai-Fun Mun, Hing-Yan Lee: Finding Interesting Patterns Using User Expectations. IEEE Trans. Knowl. Data Eng. 11(6): 817-832 (1999)

Suzuki 96 Einoshin Suzuki, Masamichi Shimura: Exceptional Knowledge Discovery in Databases Based on Information Theory. KDD 1996: 275-278

Suzuki 97 Einoshin Suzuki: Autonomous Discovery of Reliable Exception Rules. KDD 1997: 259-262

Suzuki 00 Einoshin Suzuki, Shusaku Tsumoto: Evaluating Hypothesis-Driven Exception-Rule Discovery with Medical Data Sets. PAKDD 2000: 208-211

Hoschka 91 Peter Hoschka, Willi Klösgen: A Support System for Interpreting Statistical Data. Knowledge Discovery in Databases 1991: 325-346

Klosgen 96 Willi Klösgen: Explora: A Multipattern and Multistrategy Discovery Assistant. Advances in Knowledge Discovery and Data Mining 1996: 249-271

B. Liu 99b ing Liu, Wynne Hsu, Yiming Ma: Pruning and Summarizing the Discovered Associations. KDD 1999: 125-134

Yugami 00 Nobuhiro Yugami, Yuiko Ohta, Seishi Okamoto: Fast Discovery of Interesting Rules. PAKDD 2000: 17-28

Keywords: Parallel universe for rule discovery, Rule pair, Exception rule, Interestingness, Rule discovery, Medical application

Edge Effects in Pattern Detection

Edward A. Tricker (Imperial College London, GB)

An edge effect occurs when the region of support of a distribution ends abruptly. Pattern discovery methods based on local estimates of probability density are compromised by edge effects since apparent regions of artificially high density occur at the edges. This problem is more severe the larger the number of variables defining the data space. We analyse the problem and describe tools for overcoming it.

Parallel Universes: Multi-response Optimization

Claus Weihs (Universität Dortmund, D)

One universe may be seen as one view of the universe or as one quality of the universe, represented by one response variable. Parallel universes may be seen as many views of the universe, as many qualities or many criteria represented by multiple responses.

The talk will consider the case when all responses are optimized in terms of (at least in principle) the same set of influential factors. If optima are inconsistent, i.e. in different localities, how to resolve the conflict? The talk will discuss comprises: Desirability indices for the reduction to one dimension, Pareto fronts and sets for the multivariate perspective, and desirability functions for an a-priori restriction of Pareto front and set to desired parts.

Keywords: Multi-response optimization, Pareto front, Pareto set, desirability function, desirability index

Joint work of: Weihs, Claus; Trautmann, Heike

Fuzzy Clustering in Parallel Universes

Bernd Wiswedel (Universität Konstanz, D)

We present an unsupervised clustering method for learning in Parallel Universes based on the fuzzy c-means algorithm. The method uses fuzzy membership values to encode degrees of clustering contributions for each object to each universe. The algorithm iteratively learns these newly introduced values and - similar to the classical c-means - partitioning values in each universe. The outcome of the algorithm are clusters spread across different parallel universes, each cluster modeling only a small subset of the data. We will furthermore discuss two extensions to this basic algorithm to overcome the negative influence of noise and/or outliers (caused by the partitioning property of the underlying algorithm) and the problem of specifying the number of clusters in advance. Firstly, we will show how incorporating an additional term to the global objective reduces the influence of noise. This term represents a virtual noise universe with one single cluster, attracting objects that do not contribute to any of the clusters and hence reducing their impact on the cluster formation. The second extension discards the side constraint requiring that universe memberships need to sum to one, allowing potential overlaps of clusters (in different universes) or empty clusters.

Full Paper:
http://www.inf.uni-konstanz.de/bioml2/publications/Papers2007/WiBe07%5ffcum%5fijar.pdf

See also: Bernd Wiswedel, Michael R. Berthold, Fuzzy Clustering in Parallel Universes , International Journal of Approximate Reasoning, vol. 45, no. 3, pp. 439-454, 2007