STARTPAGE

HUMAN RESOURCES AND MOBILITY (HRM) ACTIVITIES

MARIE CURIE ACTIONS
Research Training Networks (RTNs)

PART B

PASCAL-RTN

Marie Curie Research Training Network in

Pattern Analysis, Statistical Modelling and Computational Learning

April 2, 2003

Abstract

PASCAL-RTN is a four year interdisciplinary Marie Curie Research Training Network designed to complement and consolidate the training and community-building activities of the proposed FP6 Network of Excellence PASCAL. The theme of both networks is the confluence of pattern analysis, statistical modelling, computational learning and large-scale optimization to create new and very powerful methods for analyzing large quantities of sensed or warehoused data. Their approach is rooted in rigorous statistical learning and optimization theory, and they will cover learning and pattern recognition aspects of a range of important sensor modalities, including: natural language understanding; speech analysis; text/web mining; computer vision based image and video understanding; and user interface topics including haptics, novel interfaces such as brain computer interaction, user modelling, and data fusion for multimodal interfaces. PASCAL aims to build a vibrant interdisciplinary European research community centred on rigorous and novel applications f statistical pattern analysis and machine learning. PASCAL-RTN will consolidate this effort, providing a core of top-ranking young researchers with in-depth training in this emerging discipline, and further breaking down interdisciplinary barriers by allowing focused research attacks on key 'watershed' issues to be funded. Topics covered will include: computational learning theory for adaptive systems and for problems with many overlapping classes; the relationship between Bayesian and frequentist learning; structured models; optimization methods for large-scale statistical modelling and learning problems; advanced kernel based methods for visual object detection, speech analysis, and text/web mining; and learning based approaches to user modelling and data fusion.

B1. Scientific Quality of the Project

B1.1. Research Topic

We live in an age where rapid and reliable access to information is the key to success, and where social cohesion requires that all users, including non-technical ones, have sufficient access. Conversely, we are surrounded by raw data sources of all types, from traditional senses such as vision, touch and hearing, through to huge new data repositories such as digital libraries and the World Wide Web. Almost all of these sources are unstructured in the sense that extracting the needed information from them requires highly nontrivial search and interpretation capabilities and constant adaptation to change. This is true even for nominally structured sources such as the Semantic Web, because usage and access patterns vary over time, and because suppliers and consumers are individuals with different needs and capabilities and often with conflicting goals. But it is doubly true both for raw sensed data, which is naturally unstructured, redundant and somewhat chaotic, and for modalities such as text, which is optimized for the high-level cognitive abilities of humans rather than the lower-level ones of current automated text understanding systems.

Hence, there is a pressing need for systems that can efficiently filter, merge and interpret information from disparate sources, and that can easily adapt to changing access conditions and different user's information needs. To be useful, such systems must scale to the huge size and wide range of categories of current digital repositories, and to be reliable they must be based on well-founded theoretical and algorithmic analysis. Achieving this goal will require the combined expertise of a number of domains, including: machine learning, statistical pattern analysis and random processes for the basic underlying models and methods; large-scale mathematical programming for database-scale learning algorithms; and application expertise from vision, speech, and text/web processing, and from user interfaces and user modelling.

Our proposed FP6 Network of Excellence PASCAL (Pattern Analysis, Statistical Modelling and Computational Learning) is designed to achieve this integration, building a coherent, vibrant European research community centred on machine learning, pattern recognition and their applications from the fragmented subcommunities that exist at present.

The current application, PASCAL-RTN, is for the Research Training Network component of PASCAL. PASCAL-RTN is an integral part of our overall community-building strategy. It will break down interdisciplinary barriers and increase critical mass by allowing focused research efforts on selected key 'watershed' issues that currently either separate communities, or that are not receiving the attention they deserve owing to their interdisciplinary nature. Topics to be covered will include: computational learning theory for adaptive systems and for problems with many overlapping classes; the relationship between Bayesian and frequentist learning; structured models; optimization methods for large-scale statistical modelling and learning problems; advanced kernel based methods for visual object detection, speech analysis, and text/web mining; and learning based approaches to user modelling and data fusion.

Like PASCAL, PASCAL-RTN is a deeply interdisciplinary network combining groups working on machine learning, statistical pattern recognition, random processes, mathematical programming, computer vision, speech understanding, natural language and text/web processing, and user interfaces.

B1.2. Project Objectives [1]

The overall objective of both PASCAL and PASCAL-RTN is to weld the widely scattered subcommunities working on computational learning and statistical pattern recognition methods in Europe into a single active community focused on well-founded methods and algorithms, and to encourage the wide diffusion and uptake of this emerging technology with a coordinated program of focused research and pump-priming activities, curriculum development, workshops, summer schools and visitor programs, challenge problems and datasets, reference algorithms, and an expertise brokerage and virtual science park program. The particular role of PASCAL-RTN in this will be to further break down barriers by

[1] This and some of the subsequent sections are somewhat longer than recommended, but we feel that this is reasonable given the comparatively large size of the requested network, and its "precision targeted" organization, *i.e.* the fact that is designed to rapidly break down interdisciplinary barriers by applying localized research pressure at a relatively large number of separate, carefully-selected strategic 'weak points' on domain boundaries. The full proposition is well under the total maximum size implied by the sum of the page limits, which is 29 pages plus 2 pages per team.

providing an opportunity for focused research on a range of key interdisciplinary topics, while also providing a core of talented young researchers with both broad and in-depth interdisciplinary training in the PASCAL subject area. The training will cover computational learning, mainstream and Bayesian statistics, optimization methods, and application topics including vision, speech, text/web, haptics, user interfaces and user modelling.

Particular scientific focuses for research in PASCAL-RTN will include:

Kernel methods for clustering data *(RHU London, 36 months postdoc).* There has been some work in developing kernel based methods for clustering. This has aimed at applying two well-known approaches to clustering in kernel defined feature spaces, hence extending the applicability of the methods to higher dimensional and more powerful representations. One example of this development has been the implementation of the k-means algorithm in kernel defined spaces, while another has been the extension of spectral clustering methods to such spaces. In both cases the developments have been fairly routine in that they have not made any adaptations of the basic algorithms to take advantage or avoid the dangers implicit in the high dimensional spaces. This is in contrast to the extension of for example linear classification to kernel defined features spaces, where the margin maximization constraint is introduced in support vector machines to overcome the danger of overfitting inherent in the large numbers of dimensions.

The challenge that we propose to tackle is to develop clustering methods that are well-founded in a statistical sense and hence that can be designed to guard against the dangers of high dimensionality, while still reaping its benefits. If successful we believe that this could deliver an advance in clustering technology akin to that evidenced by support vector machines for the task of classification.

Visual recognition with thousands of overlapping classes *(INRIA Grenoble, 36 months PhD plus 6 months postdoc).* Humans appear to be capable of recognizing on the order of 100,000 visual object classes, and human-constructed visual categories are at once overlapping & highly structured — objects can be classified according to different criteria at different levels of specificity, and images typically contain several interacting objects — and very fluid — ad hoc classes can easily be constructed on-the-fly for any given application. Most current machine learning techniques derive from two class, yes/no classifiers, and although multi-class extensions exist, it is difficult to see how they could be extended to cover 100,000 complex, highly structured, overlapping, changeable classes. A rethinking of the basic theory and some new algorithmic ideas are needed to reach this level of generality, and this will require a close dialog between machine learning and vision researchers. A possible approach would be to develop multicategory learning methods based on a dynamic Bayesian network like representation of the lattice of classes, with implicit generative models giving dynamical class-activation probabilities at the nodes and probability contrast rules (perhaps conventional classifiers based on specialized features extracted on-the-fly) representing specialization/discrimination rules on downwards arcs and generalization rules on upwards ones. But it is unclear how to ensure good generalization in such a framework, and good localization / pruning methods would needed to be developed to keep the calculations in such an approach tractable.

Analyzing approximate inference methods for Bayesian networks *(NCRG Aston, 24 months postdoc).* Bayesian networks provide a powerful framework for modeling statistical dependencies between data. The price that a modeler has to pay for the high degree of flexibility of these models is the vast increase in computational complexity when the number of nodes in the network is huge. Hence, a variety of approximate inference techniques for Bayes networks (partly motivated by ideas of statistical physics) have been developed which seem to combine fairly accurate predictions with manageable computational costs. However, a major drawback of approximate approaches is the fact that the accuracy of the approximation is usually unknown. Probabilities computed from an approximate inference algorithm may differ significantly and in an uncontrolled way from the "ideal" model probabilities and also from the true (but unknown) probability of the data. Further, it is often unclear how to improve these approximations systematically. The proposed project aims at developing a sound theoretical basis for assessing and improving the quality of practically relevant approximate inference methods for Bayes networks.

Speech modelling for advanced telephony *(LIA Avignon, 24 months PhD, 24 months postdoc).* Research for new and more advanced telephone services is an important component in the framework of ambient intelligence. Processing telephone speech should be carried out with robust recognizers capable

of dealing with the bandwidth, types of noise and distortions which are present in telephone speech. Although great progress has been observed in recent years and very complex systems have been effectively deployed thanks to the results of the previously funded IST SMADA project, there is ample margin for improvement. Among the open problems, attention will be focused on using language and semantic models for assessing the confidence of computer interpretations of segments of telephone dialogues.

Randomized algorithms plus **Learning in data streams** *(UPC Barcelona, 12 months PhD plus 12 months postdoc (year 3))*.

1. Randomized algorithms for large-scale use of support vector machines. The impressive performance of support vector machines for prediction problems is counterbalanced in practice by their relatively high computational cost. Recently, abstract randomized algorithms for training support vector machines have been proposed whose theoretical cost is quasilinear in the number of training points (as opposed to the quadratic as in the usual approaches). However, these algorithms are still far from practicality owing to the more complex implementations required and their strong dependency on the dimension. The challenge is to transform these randomized algorithms into ones that can can be used in practice to solve medium or large scale prediction problems and, at the same time, that can be theoretically validated.

2. Learning in time-varying data streams. The *data stream* is a model of the situation when a program has access to a rapidly arriving, potentially infinite, stream of data in which each item is seen only once. Algorithmics for data streams is now the subject of intense research efforts, but the results on learning from data streams are relatively scarce. We intend to develop systematic ways of designing learning algorithms working on data streams. These should be incremental and able to deal with time-varying streams, *i.e.*, able to change the learned hypothesis as the stream statistics change. Hence this challenge is an instance of that of developing a theory for adaptive learning systems. An important tool will be recent probabilistic techniques known as *adaptive sampling*, useful for scaling up data mining algorithms to large datasets.

Extended multi-class learning *(IDA Fraunhofer, 40 months postdoc)*. A common property of many multi-class problems is that there are relationships between the classes that — if discovered — could greatly simplify the learning task. To do this one either needs to supply additional class information, or this information needs to be extracted automatically from the training data. There are many unsolved problems of how to do this and how to guarantee good performance of learning algorithms. New formulations of this will most likely lead to new optimization problems that need to be solved in order to solve a given learning problem, so part of the challenge will also be to design and implement efficient algorithms for solving the optimization problems for large scale learning problems.

Optimization methods for large scale learning *(U Bayreuth, 36 months PhD)*. The nonlinear structure of modern mathematical learning models, for example methods based on nonlinear kernels, and the large, often huge, scale of the resulting optimization problems, require the development of new algorithms and the careful implementation of highly efficient and robust codes. Of particular interest will be so-called Sequential Convex Programming (SCP) methods for general non-convex large scale nonlinear programming. The strictly convex and separable subproblems of the same size can be solved efficiently by interior point methods, where the size of systems of linear equations to be solved can be reduced to the expected number of active constraints and where special sparsity patterns can be exploited. The SCP approach will allow new separating surfaces, *e.g.*, for general support vector machines, to be constructed using convex hyperplanes. Since the methods are designed to solve large scale non-convex programs, the resulting algorithms and codes are applicable in all other topics of the project where mathematical optimization is involved.

Sequential Querying and Experimental Design *(EURANDOM, 24 months postdoc)*. The concept of sequential querying refers to the ability to learn between each observational step in order to direct the next query or interrogation of the real world or data base. This idea is not new having its pedigree in sequential experimental design and similar areas, such as sequential search. The naive version is that the next query should be taken so that the expected increase in information on receipt of the new data is optimal. More sophisticated methods should use full scale dynamic programming (Bellman). In this case, one would take into account at time t that one was behaving optimally at time $t+1$, and so on up to a fixed time horizon, T. Dynamic programming would work back from the time T to the present. The

scientific challenge is both theoretical — to integrate these ideas, particularly dynamic programming, into the computational learning environment — and practical — namely to construct fast algorithms that keep up adequately with the data flow of the application. Applications are to computer vision and imaging, directional microphones, and indeed any area where gains are to be made from directing attention, in the sense of cognitive science. The project will work closely, therefore, with application areas via multi-disciplinary teams to make the project more applicable and feasible.

Biologically motivated computational learning principles *(TU Graz, 36 months PhD)*. Recent results in neuroscience have shown that learning in biological systems uses a number of learning principles (such as spike time dependent plasticity in combination with context-dependent learning rates and reward-related consolidation) that are not represented in the current repertoire of machine learning algorithms. The scientific challenge is to build computer and theoretical models of those among these learning principles that have a good chance of being portable to machine learning, and to compare them with the performance of other machine learning algorithms. We expect that some of these new learning principles will exhibit superior performance in the areas of novelty detection, fast scene analysis and on-line learning. In order to test this we will implement them in a mini-robot (khepera).

Information-theoretic learning of Bayesian networks and structured models *(CoSCo Helsinki, 36 months PhD plus 12 months postdoc)*. The objective of this research is to develop and study the Minimum Description Length (MDL) approach to modelling and its relationship to other probabilistic approaches used in computer science and statistics. Formally, MDL modelling can be approached from various directions including Fisher's Maximum Likelihood Principle and minimax formulations leading to universal coding which extend Shannon's Source Coding Theorem. In particular, research is focusing on certain recent formulations of MDL (NML, Kolmogorov sufficient statistics), and also on its predictive form PMDL, the properties of which are not yet well understood. The emphasis on applying the theoretical work will be in computationally efficient model selection and prediction problems in computer science, but the results are widely applicable in statistical modelling in general, including regression, cluster analysis and curve estimation.

Complex categories for text/web analysis *(MU Leoben, 36 months PhD)*. In the area of large-scale text or web analysis new well-founded methods are needed for dealing with very complex class hierarchies. In such applications typically hundreds or even thousands of classes organized in deep hierarchies need to be categorized. Flat categorization of classes without using the hierarchy is relatively well understood, but new approaches are needed to significantly improve categorization by using hierarchical structure. Such a new approach will need to combine methods from information retrieval, machine learning, statistics, optimization and natural language processing. It can be anticipated that a combined effort will significantly improve the accuracy of such retrieval, filtering, and classification systems and make them much more accessible to a wide range of users.

Combining text-mining with link analysis *(JSI Ljubljana, 12 months PhD plus 20 months postdoc)*. The idea of this research is to combine methods coming from two areas of analyzing unstructured data: textual data (represented by areas such as text-mining, information retrieval, natural language processing) and network structure (represented by areas such as social network analysis, link analysis, graph theory). There are a number of still unexplored research possibilities on how to enrich textual representation via network/graph representation. The goal of this research is to connect knowledge from both areas.

Learning complex probability models for video object detection *(RRG Oxford, 36 months PhD plus 6 months postdoc)*. The challenge is to annotate video with the objects it contains. By an object we mean a visually defined object class such as a building, person, animal, *etc*. Such annotation is difficult because of the usual problems that computer vision faces: an image is affected by viewpoint, lighting and occlusion, all of which are unknown; but also by variability within an object class, for example the variability in the appearance of a human arising from different positions of the limbs through to different facial expressions. To surmount these difficulties requires elements from both computer vision and learning: computer vision to account for the viewpoint, lighting etc, and learning to account for the variability within an object class.

In particular, we will investigate fitting two types of parametrized probabilistic models in order to recognize visual object classes. The first is for classes that contain repetition, *e.g*. buildings (churches, skyscrapers, houses), natural scenes such as forests, mountains, waterfalls. We will use groupings to find

these repetitions (computer vision part) then learn which are relevant over training sets (learning part). The second model is for articulated objects such as humans. Here the model covers the articulated structure and its visual shape (computer vision part) but its components and the parameters of the articulation must be learned from training examples — *e.g.* video (learning part).

Learning structured language models for speech recognition *(UJM Saint-Etienne, 12 months postdoc)*. Language models are an essential part of speech recognition applications. Current models are very simple and do not take into account phrase structure or long term dependencies. Work on the learning of more complex models (based on graph or automata type of structures) is promising, but such models still require a lot more work in order to compete with simpler alternatives. The scientific challenge is to provide learned automata with up-to-date techniques (typing, use of background knowledge, enhanced parsing, theoretically based smoothing...) to beat state-of-the art language models (n-grams).

Sequential learning, adaptation and optimization *(I3S Sophia Antipolis, 36 months PhD)*. Most learning tasks have a more precise objective than just learning, with optimization as a typical example. We thus concentrate on the problem of optimizing an unknown function (which may describe the performance of a system), with respect to some user chosen input variables. The function is learned "in real-time", that is, at the same time as the search for its optimum. To each set of input variables that is proposed, the system responds by a noisy realization of the function value. For instance, standard adaptive control corresponds to the situation where the systems contains some dynamics, and a parametric form of the function is known (derived from a parametric model of the system itself). Applications in medical trials concern the situation where the realization observed for a choice of inputs (a pharmaceutical dose) is binary (success or failure), and the function to be optimized describes the probability of success. To determine optimal inputs, one needs to learn the form of the function to be optimized, and at the same time learning is made from the inputs that are proposed. The challenge is to learn as quickly as possible (approaches of the Kiefer-Wolfowitz type are extremely slow), while proposing inputs as close as possible to the optimum value, the two objectives being clearly contradictory (hence the name "dual control" in control theory). Further advances will require combined efforts in several directions, including control theory, statistical modelling and inference (sequential methods in particular), and non-parametric modelling.

Recognition of complex visual events *(KTH Stockholm, 33 months postdoc)*. Recognition of dynamic events such as human actions is complex because of large variability in appearance due to variation in actor, clothing, viewpoint and illumination etc. The challenge is to derive descriptors and representations that are invariant to these conditions as far as possible. These invariants have to be complemented by the learning of priors based on evaluation of large data sets of human action. The systematic generation and representation of these priors in terms of manifolds in representational space is the main scientific challenge.

Computational learning of structured mappings *(MPI Tübingen, 36 months PhD, 12 months postdoc)*.

1. Learning to predict structured outputs with applications to text problems (PhD). The goal of this challenge would be to design efficient algorithms to learn a mapping from a structured input (like a text) to a structured output (like another text). This problem cannot be addressed by current classification, regression or multi-valued regression techniques. The idea would be to use a kernel to measure similarities of output texts and to create an output feature space (this is an idea recently proposed by Weston, Schölkopf, Chapelle, Elisseeff and Vapnik, at NIPS 2002). A major application of this technique would be the design of a question-answering system, able to answer queries on a certain topic based on a FAQ database. Other applications could be found in various text processing problems.

2. Incorporating prior knowledge in the pre-image problem (postdoc). Related to the above challenge of learning to predict structured outputs comes an important issue: the pre-image problem. Indeed, to predict structured outputs, one first builds a mapping into a feature space for outputs, and then has to recover the correct output from the predicted feature values. This can be computationally hard (it may involve a search across the whole space of possible outputs). One way to speed up the computations would be to make use of prior knowledge of likely outputs to introduce constraints on the search. The challenge is to identify appropriate constraints (like properties of the structured outputs, invariances,...) and to design the corresponding algorithms.

Learning Bayesian networks with latent structure of known type *(U Aalborg, 12 months postdoc)*. Methods should be developed to flexibly specify and efficiently learn Bayesian network structures with unobservable variables, exploiting additional knowledge about their type, for example that the networks are genetic pedigrees. This type of learning is relevant for complex body identification problems such as occur in connection with major disasters.

Statistical learning in a mixed-source setting *(UC London, 24 months postdoc)*. The general problem of statistical learning and pattern recognition has in the recent two decades received a significant theoretical advancement through the introduction of the Vapnik-Chervonenkis theory also known as statistical learning theory which subsumes the Probably Approximately Correct model of learning. This theory represents a major stream in the field of Computational Learning and is focused on the representation and analysis of sample-based learning, namely, explaining the performance and limitations posed on algorithms that learn by examples. However, behind any learning machine stands a major hidden influence that takes the form of prior assumptions or side information that is implicitly based on the choice of models used, the choice of loss criterion, and the algorithmic hypothesis search. In order to develop better learning algorithms it is crucial to understand the influence of such different sources of information on the accuracy of sample-based learning. This proposal aims to investigate various learning problems under such mixed-source settings, and analyze the error rates as a function of sample size and type and amount of information.

A framework for kernel-based on-line learning *(U Milan, 12 months postdoc)*. In many learning applications, ranging from adaptive interfaces to document filtering, the algorithm must operate in real-time on a stream of incoming data. On-line learning algorithms work by making incremental adjustments to a current model (*e.g.*, a generalized linear model), and thus are well suited for stream-based learning problems. Research in on-line learning is deeply connected to statistical learning as well as to game theory. Many learning algorithms have specific geometrical properties (rotational invariance) allowing them to be run efficiently in large and complex feature spaces such as reproducing kernel Hilbert spaces. The objective of this research is to investigate a general framework where kernel-based on-line algorithms can be designed and their learning performance rigorously analyzed.

Machine learning for structured textual information retrieval plus **Reverse modeling of dynamical systems** *(UPMC Paris, 52 months PhD)*.

1. Machine Learning for structured textual information retrieval. The project concerns machine learning for information retrieval in structured document corpuses, particularly XML. The concept of 'document', whether textual or multimedia, has considerably changed with new norms and formats of structured documents. This is due to the development of the web and to new access needs for multimedia and textual contents of documents. Standardization organizations have proposed, and are currently concerned with the defining, new document models that allow documents to be represented in a richer way (*e.g*. Smile, RDF). Such standards include not only raw information but also structural information (as in DTD or XML) and metadata. The development of search engines able to simultaneously handle all of these forms of information is an important challenge in both the industrial and research fields. It requires the development of tools allowing corpuses of structured documents to be both stored and retrieved in a flexible way. This is a new and challenging task. In 2002, there was a first attempt to build an evaluation corpus for structured document search engines, to define fundamental functionalities of these engines and to perform preliminary tests. The Computer Science Lab of UPMC was one of the 30 international teams in this meeting. This meeting will be held once again this year and following this, will be integrated into the NIST competitions.

2. Reverse modeling of dynamical systems. Many real-life time series, such as financial time series or gene expression profiles, show a stochastic behavior governed by a hidden process. Analysis of such time series is thus a challenging task, knowing that both the extraction of the features of the driving process and the understanding of the observed signal characteristics (with the goal of predicting future time series values) are key issues in many research fields. In this framework, Bayesian dynamical methods appear to be principled and powerful tools to deal with these issues. Although some dynamical models are well studied, much still has to be done in the context of introducing and fusing expert knowledge, Bayesian network structure selection, mixing of models, and integration of kernel methods.

Cognitive 3D reconstruction *(ETH Zurich, 36 months PhD).*

3D visual reconstruction algorithms still frequently fail to produce convincing reconstructions of scenes, even ones with quite a simple structure. In this project we want to combine scene categorization with 3D reconstruction. By recognizing that a scene belongs to a certain category (*e.g.* a room, a building, ...) the reconstruction can be enhanced. On the other hand, the availability of 3D data can ease the categorization of the scene. In particular, we aim at automatically producing convincing scene models from multiple, uncalibrated photos. The 3D reconstruction pipeline already exists, but in this project we want to add a cognitive layer as just described. We plan to tackle the problem of feature selection in high-dimensional spaces (e.g. based on large sets of wavelet or similar filters), by estimating the decision boundary feature matrix using the gradient of the SVM decision functional evaluated at the SVs. Moreover, the intention is to study feature hierarchies and cascades of classifiers in order to arrive at the necessary combination of speed and performance for scene categorization. The system should learn from additions and corrections interactively made by the user how to adapt 3D reconstructions of scenes belonging to the same category, so that gradually the process becomes maximally automated.

B1.3. Scientific Originality of the Project

For each research topic outlined in section B1.2, this section describes the contribution expected relative to the current international state-of-the-art in that topic.

Kernel methods for clustering data. (RHU London) The first task will be to develop a rigorous statistical theory for the generalization from a clustering found in training data to its fit with new data generated according to the same distribution. It is bounds of this type that are used to motivate the capacity control techniques used in the application of kernel methods to both regression and classification. The techniques most appropriate for this will be the Rademacher complexity bounds on loss functions, where the appropriate loss function for clustering will be a measure of misfit with the empirically determined clusters.

The second stage of the work will be the development of optimization criteria that minimize the expected loss assessed in part one. The optimization will be performed over candidate sets of clusterings that would no longer be restricted to standard sets of prototypes. For example we anticipate that a set of k one class support vector machines could provide one such candidate set.

The third stage will be the development of efficient algorithmic strategies that are efficient both in terms of rate of convergence as well as in the quality of solution obtained. One method will be generalizations of the EM algorithm used for k means to for example k one class support vector machines. Such an approach will however exhibit the difficulty of non-convex optimization criteria typical of clustering. As with soft margin support vector machines relaxations of the criteria can lead to efficiently solvable tasks. This is also the case for spectral methods of clustering. Here the research will focus on extending these methods in analogy with the use of semi-definite programming in graph bipartitioning. The link with optimization within the overall TMR will provide additional support for this phase of the work. The aim is to develop algorithmic strategies that are efficient and deliver solutions that come with guarantees of being within some fixed fraction of the optimal solution, as is the case for the graph bipartitioning.

Visual recognition with thousands of overlapping classes. (INRIA Grenoble) Current machine classification methods are largely descendants of two class discriminants, almost all assume a 'flat', rigid, pre-specified set of classes, and few scale well to very large numbers (thousands) of classes. To handle 'unengineered' data such as natural images, we need produce learning methods that can handle something like the fluid, dynamical, overlapping nature of human-created categories, while still remaining well-founded and algorithmically tractable.

Analyzing approximate inference methods for Bayesian networks. (NCRG Aston) The statistical performance of approximate inference algorithms for Bayes networks is an open challenging problem which has not been put on a firm theoretical basis. To achieve this goal, we propose a combination of ideas from statistical physics and of rigorous methods developed in the machine learning community. Concentrating on the so called TAP approximation [2] widely applicable to undirected dense graphical models, we will first aim at deriving systematic corrections based on perturbation methods of statistical

[2] see *e.g.*,"TAP Gibbs Free Energy, Belief Propagation and Sparsity", Lehel Csató, Manfred Opper and Ole Winther; Proceedings of NIPS 2001

physics. Second, the fact that the quality of the approximations is data dependent suggests the development of a *luckiness framework* providing bounds for a "probably good approximation" with respect to the distribution of the data. Reasonable candidates for luckiness functions are motivated by ideas from statistical physics. Finally, we will try to develop a *PAC-Bayes* framework allowing to get performance bounds for certain classes of Bayes networks when approximate posterior probabilities are used for prediction rather than the exact ones.

Speech modelling for advanced telephony. (LIA Avignon) Searching for concepts can be combined with searching for words. This suggests that statistical language models (LMs) could be adapted based on expectations of concepts predicted by a system belief. With this perspective, it is important to notice that, while the observation of only certain words may be sufficient for hypothesizing a conceptual structure, complete details of word phrases expressing a conceptual structure have to be known in order to adapt a generic LM to the expectation of such a structure. Furthermore, reasoning about recognition and interpretation results will produce linguistic confidence values to be used by a dialogue strategy.

Randomized algorithms plus **Learning in data streams.** (UPC Barcelona)

1. Randomized algorithms for large-scale use of support vector machines. The outcome of our work should a proof-of-principle prototype of an efficient randomized algorithm for training support vector machines. This prototype should indicate that, upon further engineering work, support vector machines could be applicable to medium or large-scale prediction problems, such as those appearing in data mining.

2. Learning in time-varying data streams. Handling time-varying sources of data is an important concern in many practical situations, which is today addressed by ad-hoc methods. A systematic way of designing learning algorithms for scenarios where the distribution or nature of the data changes over time would enlarge the applicability of many results in computational learning.

Extended multi-class learning. (IDA Fraunhofer) Learning the classification of problems with many classes often occurs in practice, e.g. in protein structure prediction, speech analysis and computer vision. Most learning algorithms do not exploit that there are relations between the classes and that in practice these relations are known or at least some additional information about their relation is known. We anticipate a great breakthrough if we find a principled way of describing and learning class relationships and would expect a large impact in many application fields when efficient algorithms designed and implemented.

Optimization methods for large scale learning. (U Bayreuth) Traditionally, modern machine learning algorithms reduce to solving large-scale linear, quadratic, or semi-definite mathematical programming problems. Optimization has thus become a crucial tool for learning, and learning a major application of optimization. On the other hand, sequential convex programming (SCP) methods for smooth large scale nonlinear programming have been invented by mechanical engineers to solve large topology optimization problems. Theses methods provide a novel technology never applied before to the problems under consideration. The special structure of the mathematical model, for example in case of a general support vector machine, provides an additional challenge to further improve efficiency and scale. New convex separating surfaces can be investigated besides the standard nonlinear kernels used until now.

Sequential Querying and Experimental Design. (EURANDOM) This project can be seen as one component of a somewhat under-researched area in computational learning, that is to say dynamical versions. Other areas of science notably control theory and (as mentioned) sequential experimental design are in advance of computational learning in being truly dynamic. However, they suffer from rather strict paradigms and do not as yet have the benefits of robustness that arise in computational learning. Expected innovations include: (*i*) the use state space methods to control the learning; (*ii*) the combination of static and new dynamic programming methods and (*iii*) the use of Shannon and heuristic search rules to optimize the querying.

Biologically motivated computational learning principles. (TU Graz) Learning principles for neural networks have so far primarily explored hypothetical neural learning principles that were based on rather old biological data. Hence the extraction of portable new learning principles from more recent neurobiological data and their evaluation in a machine learning environment is highly innovative.

Information-theoretic learning of Bayesian networks and structured models. (CoSCo Helsinki) In various fields in computer science (AI in particular) there is a need for model building from data. Similarly the purpose of statistical model building is to extract information from a given set of data. Minimal encoding approaches to modelling have their intellectual roots in the descriptive complexity theory of Kolmogorov, Chaitin and Solomonoff, which can be used to define inductive inference in general (or statistical inference in particular) as the search for the shortest program for data. By restricting the models to be probability distributions combined with deterministic equations, and defining the descriptive length in probabilistic terms, one is able to avoid the non- computability issues inherent to descriptive complexity. This has lead to a rich theory of approaches equating describing data with coding and using code lengths for model building. Because of the descriptive philosophy, such approach provides with new insights into familiar statistical procedures, and also serves as an objective platform from which we can compare for example Bayesian and non- Bayesian statistical procedures alike. The theoretical minimum-encoding formalism has also interesting connections to the on-line statistics and prequential formalisms, and better understanding of these relationships would advance research in all these areas.

Complex categories for text/web analysis. (MU Leoben) For multi-class categorization currently most systems work in "flat mode", which means that all classes are treated equally and known class dependencies are rather not taken into account. The main problem of hierarchical categorization is the quantitative evaluation of different possible classification paths through the hierarchy. Combining methods from machine learning and statistical analysis we expect new quantitative methods for path evaluation which allow hierarchical categorization as a standard tool. Possibly such new methods will combine margin bounds known from support vector machines with likelihoods from Bayesian analysis.

Combining text-mining with link analysis. (JSI Ljubljana) Originality of the scientific challenge on combining Text-Mining with Link Analysis areas stems from the fact that both areas developed in last 10–15 years completely independent and only in the recent years with the need to analyze large data sets within data mining area possibilities appeared to join both scientific fields. The biggest potential in the combination lies in the enriched textual representations offered by general network/graph structures.

Learning complex probability models for video object detection. (RRG Oxford) The promise of learnable models is that if the process can be automated, where the parameters are learned in an unsupervised manner from training images, then for the first time it will be possible to learn thousands of object classes. Unsupervised here means that it is not necessary to first normalize the image for size nor outline (segment) the object of interest in each case. Such supervision is labour intensive and cannot possibly be carried out on the millions of images that will be required to learn a reasonable number of object classes (*i.e*. in order to recognize many of the typical objects we see).

A second point of originality is that learning offers the possibility of overcoming deficiencies of feature detectors. Feature detectors (such as regions of interest, or outline curves) are never perfect. There are always problems with drop outs, false positives, false negatives, localization errors etc. By including the detected features as part of the learnt model (for example the probability that a particular feature is correctly detected) these deficiencies can be 'modelled' in advance, and thereby overcome.

A third point of novelty is using video data (rather than image databases) to provide the training data. Video effortlessly provides 1000's of examples where within class variation can be exercised. Think of a walking human or an actor changing facial expression.

A final point of novelty is that priors can be learned for the model parameters. At present priors are typically used to generate a MLE. However, the priors may also be used (by integration) to give a full Bayesian estimation. By learning priors on what makes a visual class, this holds out the possibility of learning from far fewer training examples.

Learning structured language models for speech recognition. (UJM Saint-Etienne) The speech recognition community has worked extensively on language model issues over the last 15 years. The novel approaches that we propose (grammar induction, typing, smoothing,...) could allow the grammar based language models to become preferable in a near future. This approach is based on specific work that has taken place over the past few years by the EURISE team, which consists in a set of techniques each of which gives an advantage over standard techniques, but for which integration into a single tool remains to be done.

Sequential learning, adaptation and optimization. (I3S Sophia Antipolis) Most approaches for sequential learning and adaptive optimization rely either on rather restrictive assumptions on the system (assuming for instance that a parametric form of the function is available), or use a decoupled strategy, that is learn first, and try to optimize only when the function seems sufficiently known (two stage procedures, with standard experimental design strategy at the first stage obey this rule). The difficulty in studying the convergence properties of the procedures and their performance comes from the sequential character of the decisions (the choice of input variables). The work of T.L. Lai and C.Z. Wei forms the natural background of the analysis. Recent results [L. Pronzato, Annals of Statistics, 28(6):1743-1761, 2000] concern the case of a parameterized function, with a linear regression model for the observations. Extensions to non-linear models, and then to nonparametric approaches is quite challenging. Concerning applications, the results obtained for dose-response problems where one observes binary responses, and wishes to maximize the probability of a favourable response [L. Pronzato, E. Thierry, Sequential Methods and Applications, 11(3):277-292, 2003] open the way for advanced learning strategies in medical trials. Finally, setting a classification problem as an optimization problem renders all the machinery of adaptive control and sequential experimental design available for solving adaptive classification problems.

Recognition of complex visual events. (KTH Stockholm) Human motion analysis has traditionally been based on bottom up analysis by extraction of generic descriptors. The systematic build up and use of prior information has been negligible. The extensive use of prior information will increase the domain of application of action recognition from laboratory based simple gesture recognition to the recognition of activities in video sequences recorded under general conditions.

Computational learning of structured mappings. (MPI Tübingen) It has been recently noticed, and since then heavily used, that kernels allow to treat in a unified way problems in which the input data comes in a non-vector form. Hence the development of many kernels for structured data like text or sequences. However, treating outputs in a similar way is an idea that is only emerging now and that has yet to be explored. First making this step is highly non-trivial since it involves the solution of the so-called pre-image problem. Solving this problem is likely to require a major breakthrough which might have consequences on other domains as well, since it is an instance of the so-called inverse problems (which appear in many situations such as medical functional imaging, control...). Also being able to work with structured outputs opens the way to a being able to treat a considerable number of problems that cannot be addressed at the moment. These problems include image completion, text translation, automatic chat systems,... and probably many new applications would come out as the techniques become available.

Learning Bayesian networks with latent structure of known type. (U Aalborg) Currently only ad hoc or inefficient methods exist which are not able to deal with complexity of the problems encountered in mass disasters.

Statistical learning in a mixed-source setting. (UC London) There has been some related work in the area of information theory which deals with the transmission of information through channels under some side-information which is available at the receiver end. But these are based on the Shannon theory of communication which is very limited when applied to the general problem of learning. More recently, in the area of information-based complexity, Ratsaby & Maiorov (1998) introduced a measure of functional approximation by classes of finite pseudo-dimension and used that to measure the value of partial information for learning certain specific target-function classes. The current proposal aims to investigate in a more generic approach the learning error rates for settings with mixed-sources of side-information, *e.g.*, algorithm dependent biases and information expressible in other ways.

A framework for kernel-based on-line learning. (U Milan) Kernel functions are a versatile tool, with solid mathematical foundations, for designing highly complex feature spaces that capture the application domain characteristics. Understanding the interaction between the kernel function and the learning algorithm is one of the main open problems in learning theory. Recent studies in the theory of on-line learning algorithms have produced some partial results in this direction However, a more general theory, able to connect the learning performance of a large class of on-line algorithms with the empirical behaviour of the kernel, is still missing.

Machine learning for structured textual information retrieval plus **Reverse modeling of dynamical systems.** (UPMC Paris)

1. Machine Learning for structured textual information retrieval. We will tackle this task starting with existing work in the field of information retrieval in the Computer Science Lab in UPMC. The aim of the work is to develop a formalism for information retrieval in structured documents. This should lead to the development of a search engine for XML documents. The challenge consists in the integration, in an unified formalism, of both classical database management functionalities allowing the logical structure based search, and of classical information retrieval functionalities such as plain text search. Today, only a few teams in the world propose such an integration while commercial systems are still limited.

2. Reverse modeling of dynamical systems. Up to now, Bayesian dynamical methods have mainly been studied from the algorithmic point of view and the design of efficient methods to learn their parameters. A thrilling topic not tackled throroughly so far concerns the possibility to efficiently take into account prior expert knowledge in such models. In the same time, little of the work on Bayesian networks has been devoted to the determination of their structures and new paradigms intending to do so is a recent issue. Mixing of Bayesian models fully enters this latter problem. Finally, in spite of the breakthrough that kernel methods have given rise to in the statistical learning framework, very few work evidences an effort to integrate these methods in Bayesian dynamical methods. The combination of these two kinds of methods is thus a very new, original and promising challenge. All in one, the study of dynamical Bayesian methods along the different axes we have described should provide contribution to the machine learning community both at the theoretical the applied levels.

Cognitive 3D reconstruction. (ETH Zurich) The novelty mainly lies in three aspects.

Firstly, although the use of contextual information to assist in 3D reconstruction has been experimented with before, here the system itself should provide the context through the autonomous labeling of the scene as belonging to a particular category. The models of the categories will be extracted from a supervised learning strategy.

Secondly, 2D and 3D information will have a mutual influence, as the 3D data acquired without the application of context can aid in the selection of the appropriate context. Moreover, the system should learn from interactions by a user how to best adapt 3D reconstructions autonomously.

Thirdly, we plan to experiment with novel feature selection procedures, which should reduce the dimensionality of the problem of scene categorization to a fraction of its original formulation, while keeping the performance up.

B1.4. Research Method

The focus of research in PASCAL and PASCAL-RTN will be on methods and algorithms with the following properties: (*i*) theoretical and algorithmic well-foundedness; (*ii*) efficient scalability to very large problems; (*iii*) sufficient flexibility/generality to allow customization to meet particular application needs; and (*iv*) the ability to adapt automatically to changing data source properties and user requirements.

The approach will be grounded in rigorous learning theories such as PAC- and PAC-Bayesian learning, luckiness and their extensions. Apart from providing guaranteed performance bounds, these theories have proved to be very fruitful sources of algorithmic ideas, suggesting novel learning algorithms with greatly improved performance such as Support Vector Machines. During the project, it will be necessary to extend these theories to allow adaptivity, replacing stationarity assumptions with bounded change ones.

Another aspect of our bias towards well-founded methods is that we expect many of the methods developed to be based on kernels and feature space ideas. This will allow application-specific descriptors to be coupled flexibly to powerful general learning algorithms, while maintaining good performance guarantees.

Algorithmically, modern approaches to learning transform learning problems into mathematical programming ones, thus leveraging off many years of experience in mathematical programming and powerful theoretical constructs such as Lagrangian duality. Depending on the problem and formulation, various types of mathematical programs arise, including smooth and L1 ones, definite, semi-definite and indefinite ones, and (for adaptivity) mixed discrete/continuous ones. Large-scale learning applications require the efficient solution of very large, and often exceptionally ill-conditioned, mathematical programming problems. Hence, there are rich interconnections between learning and optimization, and much work to

do here, both theoretically and algorithmically.

In terms of particular applications, there is still a significant gap between application-specific needs and the available general-purpose learning algorithms. Domain-specific feature sets and problem descriptors need to be carefully optimized for performance and recast to fit into a learning-based framework, and conversely, learning theories and algorithms often need to be adapted to fit the domain — a process that often eventually leads to generalizations being discovered and hence to an extended range of applications, rather than to specialization alone.

B1.5. Work plan

PASCAL-RTN is a four year project that will fund 13 three year PhD scholarships and 26 person years postdoctoral scholarships of between 6 months and 2 years. Most of the PhD grants are scheduled to start in the first year of funding (*i.e.* autumn 2003, finishing summer 2006), although recruiting delays are likely to even out the distribution to some extent. The postdoc grants will be more evenly spread over the 4 years.

Milestones: PASCAL-RTN scholars and their host sites will be responsible for providing short annual and final progress reports for each scholarship. The reports will summarize the research done and discuss any changes of work plan, problems or delays encountered, and opportunities that need to be capitalized on. They will also list publications accepted and submitted, and participations in training programs, workshops, software development, training visits and other PASCAL-related activities. Each PhD scholar will be expected to deliver at least one paper or technical report on the grant subject each year, or a contribution of code or other added value that the management committee judges to be equivalent. Similarly, each post-doctoral scholar will be expected to deliver a paper every 6 months. ('Delivery' means that the paper or code is to be made available on the PASCAL web server). Each scholar must also participate (and preferably, present) in at least one PASCAL workshop or summer school per year, and PhD scholars will be expected to spend a total of at least 4, and preferably 6–8, months at other PASCAL sites during their 3 year scholarship.

The individual reports, together with their lists of participations, *etc*., will be collected annually and submitted to the commission along with an annual network management report giving an overview of the overall activity and progress of the network.

Detailed annual milestones for research progress have been established for each scholarship subject, as follows:

Kernel methods for clustering data, (RHU London).
– Year 1: Scientific: the statistical analysis of clustering performance with derived criteria to be optimised. Deliverable: report summarizing the state of the art and how they meet the criteria identified by the statistical analysis
– Year 2: Scientific: Implementation of generalization of k means to new candidate sets of clusterings such as the k one class support vector machines optimizing the criteria identified in year one. Deliverable: code for the generalized algorithms with benchmarking against standard clustering methods. Report on the known methods of spectral clustering and the relationship between them.
– Year 3: Scientific: Development and implementation of spectral methods involving semi-definite programming with appropriate bounds on the efficiency and quality of clustering obtained. Deliverable: Final report including the theoretical foundations for the new spectral clustering methods and assessing their performance against the state-of-the-art.

Visual recognition with thousands of overlapping classes, (INRIA Grenoble).
– Year 1: (PhD): Report describing basic theory and approach to be taken, with simple initial experiments.
– Year 2: (PhD): Initial supervised algorithm available and paper on more refined theoretical results such as pruning.
– Year 3: (PhD): Algorithm able to handle large-scale supervised problems (1000's of classes) and paper reviewing theory and algorithms.
– Year 4: (postdoc): Paper and algorithms on unsupervised clustering / class-induction in this framework.

Analyzing approximate inference methods for Bayesian networks, (NCRG Aston).
- Year one: Development of systematic perturbative corrections to TAP approximation. Development of PAC-Bayes performance bounds and test on simple Bayes networks. Implementation in existing MATLAB code, to be made available for testing on the PASCAL web server. Publication of a tech. report giving theory and experimental results.
- Year two: Development of luckiness framework for probably good approximations, implementation of bounds in previous MATLAB code. Test on smaller networks, where exact results are feasible. Publication of results for PAC-Bayes and luckiness framework in tech. report.

Speech modelling for advanced telephony, (LIA Avignon).
– Year 1: Probabilistic models for concept hypothesis generation - prof of concept and test on a dialogue corpus.
– Year 2: Analysis of sentence and word posterior probabilities with respect to the prediction of interpretation correctness.
– Year 3: Investigation of scheduling problems for recognition and interpretation processes.
– Year 4: Production and test of a software package which generates interpretation structures from a lattice of word hypotheses, together with linguistic confidence values.

Randomized algorithms plus **Learning in data streams**, (UPC Barcelona).
– Year 1: An improved analysis of the complexity of the current randomized algorithm for training support vector machines has been completed and reported. Methods for transferring adaptive sampling techniques to the data stream model have been developed and reported.
– Year 3: The data structures for an efficient implementation of randomized algorithms for training SVMs have been identified and implemented. Theoretical work on the problems of learning on time-varying data streams has been completed, giving criteria for when such learning is possible.
– Year 4: A prototype of an efficient randomized learning algorithms for training SVMs is completed. Algorithms for some major relevant learning problems on data streams have been developed and analyzed.

Extended multi-class learning, (IDA Fraunhofer).
– Year 1: survey of literature and exploration of possibilities for modelling relations between classes. Study of existing multi-class algorithms and identification of their pitfalls.
– Year 2: development of new learning formulations (optimization problems) to incorporate and automatically extract class relations. Design and implementation of algorithms to solve the new optimization problems for large scale problems.
– Year 3: Testing on benchmark data bases and identification of application domains.
– Year 4: Application and evaluation in the identified application domain(s).

Optimization methods for large scale learning, (U Bayreuth).
– Year 1: Studying available literature, investigation of different model structures and analytical approaches, SCP methods.
– Year 2: Developing first methodology to combine general support vector machines and SCP methods, analytical investigations of mathematical structure, Lagrangian duality (of subproblem), adoption of interior point method, new convex separating hyperplanes.
– Year 3: Further analytical research, linking SCP methods and separating hyperplanes, implementation of code, collecting test examples, comparative numerical performance tests.
– Year 4: Final numerical feasibility tests, 'real life' applications, adopting code to other subjects of the project, documentation of results and codes.

Sequential Querying and Experimental Design, (EURANDOM).
– Year 1: State of the Art Report into sequential querying.
– Year 2: One academic paper submitted and one conference attended.
– Year 3: One academic paper submitted.
– Year 4: One academic paper and a conference attended.
Together the postdocs will cover the whole period of the RTN by starting second half of year 1 followed by the first half of year 2; then second half of year 3, first half of year 4. With this choice, networking and exchange of information between participants of the RTN and the Network of Excellence will be optimal.

Biologically motivated computational learning principles, (TU Graz).
– Year 1: Completion of computer model for new learning principles extracted from recent biological data, and evaluation of their performance in comparison with existing machine learning approaches.
– Year 2: Theoretical analysis of these new learning principles, implementation and evaluation in mini-robot khepera. First publication on these results.
– Year 3: Optimization of the new learning principles, and exhibition of applications domains with superior performance. Second publication on these results.

Information-theoretic learning of Bayesian networks and structured models, (CoSCo Helsinki).
– Year 1: Literature survey, formulation of the research problems to be studied in the MDL setting, scientific article on initial theoretical observations made.
– Year 2: Development of new theoretical results, results published as a scientific article.
– Year 3: Implementation of the methods developed, initial empirical tests, theory refinement, scientific article on the results found.
– Year 4: Extensive empirical validation tests with real-world datasets, two scientific articles on the results found.

Complex categories for text/web analysis, (MU Leoben).
– Year 1: evaluation of "flat mode" categorization (paper and code) evaluation of initial hierarchical categorizers (paper and code) made available to PASCAL for use in related work on text/web analysis and data mining.
– Year 2: design and test of path evaluation measures in document hierarchies (paper and code).
– Year 3: implementation and test of hierarchical classification system (paper and code).

Combining text-mining with link analysis, (JSI Ljubljana).
– Year 1: Scientific paper on one or several of the classic text analysis problems (such as classification, clustering, visualization) redefined through link analysis perspective.
– Year 2: Prototype system for implementing the combined Text-Mining & Link-Analysis technology.
– Year 3: Adding text related functionality into Social-Network-Analysis package (prototype & report).
– Year 4: Scientific paper and working prototype on combining Information Extraction technology for generation and analysis of large networks.

Learning complex probability models for video object detection, (RRG Oxford).
Each year the success will be measured by: the number of object classes that can be recognized (which should go up); and the number of training examples and degree of supervision (which should go down).
– Year 1: aim for buildings, forests, people, *i.e.* quite broad categories and not too many. Perhaps 400 training images required for each (200 contain object, 200 background, restricted range of viewpoints, *e.g.* frontal only).
– Year 2: aim for sub-classes, *e.g.* churches, skyscrapers, people-standing, people-sitting. Still 400 training images required for each (200 contain object, 200 background, less restricted range of viewpoints include foreshortened images).
– Year 3: increase of categories by order of magnitude, *i.e.* 100. No restriction on viewpoint.
– Year 4: increase by further order of magnitude, *i.e.* 1000, and each learnable from a handful of images, or a single video shot.

Learning structured language models for speech recognition, (UJM Saint-Etienne).
During year 1, the EURISE team will continue to work on the different approaches to language model learning. The post-doc will begin in year 2, so at the beginning of his/her stay the post-doc researcher will have access to all of the results and code that has been produced. A first deliverable 3 months after this would consist in a report describing the various elements that have been developed at EURISE for language model construction. The following 9 months would be spent on integrating the approaches into a unique software that would be the final deliverable. The scientific milestone corresponding to the end of the work would be a scientific paper describing results on language models using grammar induction and associated techniques that could better state of the art (n-grams) results.

Sequential learning, adaptation and optimization, (I3S Sophia Antipolis).
– Year 1: Report on state-of-the-art in parametric adaptive optimization. Implementation of algorithms

for parametric adaptive optimization (linear and non-linear observations). Report 1 on parametric adaptive optimization with non-linear observations (convergence properties of the methods, performance).
– Year 2: Conference paper based on Report 1. Report 2 on nonparametric adaptive optimization (state-of-the-art, propositions for sequential strategies). Implementation of algorithms for nonparametric adaptive optimization. Conference paper based on Report 2.
– Year 3: Depending on the value of the results obtained, journal paper based on reports 1 & 2. Application to adaptive classification examples. Report and algorithms.

Recognition of complex visual events, (KTH Stockholm).
– Year 1: definition of representations of action in terms of key frames generation of large data sets for learning.
– Year 2: recognition of general action for specific person or specific clothing.
– Year 3: recognition of general action for general person.

Computational learning of structured mappings, (MPI Tübingen).
PhD:
– Year 1: Survey of literature, acquaintance and experiments with existing KDE algorithm, identification of pitfalls of this algorithm
– Year 2: Kernel design: designing one or several kernels for treating the text queries and the text answers in a question-answering system (potentially on a restricted domain)
– Year 3: Algorithm design: designing and testing a full learning algorithm including a pre-image algorithm.
Postdoc:
– Year 1: producing an algorithm that incorporates invariances in the pre-image problem. Exploration of other types of prior constraints that can be incorporated, designing the corresponding algorithm, testing on appropriate databases.

Learning Bayesian networks with latent structure of known type, (U Aalborg).
– Year 2: Report describing the problem and initial attempts for their solution.
– Year 3: Report extending the solution to more general cases and guidance for implementation.

Statistical learning in a mixed-source setting, (UC London).
– Year 1: Study and analyze learning settings and algorithms where side-information is available in various forms about the target classifier to be learned. Conduct empirical simulations for learning with side-information. Obtain empirical learning error and generalization error-rates via cross validation.
– Year 2: Theoretical analysis of complexity of learning given side-information, computation of the sample complexity for variety of learning settings and algorithms under different types side-information. Write two papers to present at a conference and in a research journal, write a report for the continuation of this research in future projects

A framework for kernel-based on-line learning, (U Milan).
– Year 1: end-of-the-year report describing the research results, one or more conference papers, report of simulations on real-world data.

Machine learning for structured textual information retrieval plus **Reverse modeling of dynamical systems**, (UPMC Paris).
1. Machine Learning for structured textual information retrieval.
– Year 1: Bibliographical work. Familiarization with information retrieval techniques used in the systems developed in our team. This will lead to technical reports on machine learning techniques for information retrieval and on standard database management techniques. This latter work will be done in collaboration with the database team at UPMC.
– Year 2: Modeling of documents and queries using graphical models. Such models should allow expressing structural and semantic information in documents, at various granularity levels. Hence it allows multiple answering (at different levels) to a query.
– Year 3: Content and structure indexation. Develop Inference algorithms that allow reasonable answering time according to operating constraints for large databases. Learning of the models.
In the last two years, the work in progress will lead to at least one scientific paper per year related to the main objectives.

2. Reverse modeling of dynamical systems.
– Year 1: A bibliographical technical report on Bayesian dynamical methods comprising models scoring techniques. A technical report on the use of dynamical Bayesian networks for gene network inference from expression profiles.
– Year 2: A technical report on the possible methods to integrate various prior knowledge into dynamical Bayesian models. A technical report on the use of mixture of dynamical Bayesian models for gene data analysis. Source code allowing to perform mixture of dynamical Bayesian models.
– Year 3: A technical report on the different ways to integrate kernel methods into dynamical Bayesian methods and the theoretical relations with statistical learning theory. Finished PhD thesis.

Cognitive 3D reconstruction, (ETH Zurich).
– Year 1: literature review of existing, context-based 3D reconstruction methods. First version of scene categorization ready, based on photos only.
– Year 2: upgraded scene categorization algorithm, combining information from photos and the 3D reconstruction that can be produced from them.
– Year 3: algorithm for the enhancement of the 3D reconstruction based on automatically recognized contexts.

Subcontracting and New Partners

No part of the work is foreseen to be subcontracted to a third party outside the PASCAL consortium, and no research organizations from outside the consortium have a planned involvement in the contract. However, given their relatively long life span and their goal of covering, unifying and actively *reshaping* their chosen research field, it is very desirable to allow networks room for flexibility and proactivity on such issues. In particular, the NoE PASCAL includes partners who are not present in PASCAL-RTN, and it may well add others during its course. If it became clearly desirable to include such a partner in PASCAL-RTN— perhaps because a key worker or research group had moved to a different site or a newly created institute, or because the site became the major European player in the topic being researched — we would seek a means to do so, either by subcontracting part of the work, or more probably by inviting the new site to become a full partner. To allow such activities to be supported, we have asked for a small number of additional person months (10% of the total) beyond the projects listed above.

B2. Training and/or Transfer of Knowledge Activities

B2.1. Content and Quality of the Training and Transfer of Knowledge Program

The PASCAL-RTN Research Training Network will provide interdisciplinary training and knowledge transfer of exceptionally high quality by capitalizing on the training and transfer activities of the associated Network of Excellence PASCAL. PASCAL and PASCAL-RTN involve many of the top-ranking European groups working on machine learning, statistical pattern recognition, mathematical programming, and their applications. PASCAL-RTN PhD scholars will be considered to be equivalent to the PhD students originally named under PASCAL in the sense that they will be expected to participate in PASCAL at least as actively as such students[3].

The PASCAL activities most relevant to PASCAL-RTN include:

- A summer school will be held at least annually, and possibly once every 6 months, alternating between Tübingen and Canberra. This will be an intensive course that is designed to fit as much as possible of the standard machine learning curriculum into two weeks. It will be taught by a pool of leading machine learning researchers.
- The development of a PASCAL PhD accreditation program. This will build on the work of the summer school, developing a set of criteria that can be used to accredit well-designed PhD courses in this area. Initially this will be an honorary accolade, but we hope that it will become a more formal one. All PASCAL-RTN PhD scholars will be required to achieve PASCAL accreditation when it becomes available. There will also be a PASCAL curriculum development program for undergraduate and masters level courses, and PASCAL-RTN scholars will be encouraged to give feedback on this.
- A tutorial workshop program, designed to provide tutorial introductions to emerging and topical subjects.
- Interdisciplinary research workshops on specific themes or problems.
- A challenges program, designed to encourage research and competition by issuing formal challenge problems (including training and test data and performance criteria) on topical issues.
- A visitor program designed to promote interchange and transfer by allowing researchers to make relatively extended stays at other sites.
- A central web site providing access to literature, tutorials and state-of-the-art software.
- Thematic programs covering each of the main subject areas of PASCAL (machine learning, statistics, optimization, applications including vision, speech, text/web...)
- Early contact with industrial partners and start-ups that may be interested in hiring finishing scholars or exploiting their results. PASCAL will also have a consultancy-enabling program and a virtual science park designed to improve technology transfer to industry.

Early stage researchers. A typical plan of study for a three year PASCAL-RTN PhD ('early stage') scholar will be as follows: (*i*) To provide a broad basis for future training, each scholar will be expected to participate in and be ratified by a PASCAL summer school during his/her first year, in addition to any technical and career development courses required by his/her host organization and degree course. (*ii*) In subsequent years, the scholar will be expected to attend, and preferably present at, at least one PASCAL workshop, and to take any specialist courses or summer schools deemed necessary for his/her program of research. (*iii*) Scholars will also be expected to spend at least 4, and preferably 6-8, months collaborating at other PASCAL site(s) during the course of their 3 year scholarships. (*iv*) Each scholar will also be assigned a mentor, a senior researcher from another site who will be responsible for monitoring

[3] The 4800 euro/year contribution towards training, networking and transfer of knowledge supplied for each RTN scholar will be considered to be the equivalent of the 4000 euro/year that the NoE receives for each named PhD student. If a student named on PASCAL later receives a PASCAL-RTN scholarship, only one of these two allowances will be claimed.

overall progress, ensuring that the scholar plays an active role in PASCAL and undertakes a well-rounded study and career development plan appropriate for PASCAL-RTN's overall goals, and mediating in the case of difficulties. In addition to the scholar's local supervisor, the mentor will have to approve the scholar's annual report, including any requested changes of direction or workplan. (*v*) Finally, to give them early experience of collaborative research management, scholars will be expected to spend a small fraction of their time (typically 5%, maximum 10%) aiding more senior researchers with PASCAL and PASCAL-RTN related research management activities such as workshop organization, reporting, web site maintenance, *etc*.

Experienced researchers. Similar principles will apply to PASCAL-RTN postdoctoral ('experienced') scholars, but in this case a personal training, study and career development plan will be agreed between the scholar, her/his appointed mentor, and the host group. Postdoctoral scholars will be required to play an active role in PASCAL appropriate to their career stage. Apart from research, this may involve, *e.g.*, giving tutorials in their areas of expertise at PASCAL workshops, collaborative visits to other sites (appropriate to the length and goals of their scholarship), and partial responsibility for some scientific or organizational activities (under the supervision of a senior researcher).

B2.2. Impact of the Training and/or Transfer of Knowledge Program

Machine learning and pattern analysis are key strategic areas that are rapidly becoming economically important owing to the emergence of the information based economy and the growing need to manage complexity, mine huge quantities of data and build more intuitive interfaces. To some extent they are traditional areas, but they have changed enormously over the last 10 years with the introduction of: rigorous 'Computational Learning Theory' style analysis; margin based cost functions and mathematical programming based algorithms; innovative adaptive techniques such as boosting; and the extensive use of kernels to encapsulate domain knowledge about relevant descriptors, thus allowing powerful general learning methods to be applied very effectively to complex applications. These developments have caused significant fragmentation in a domain that was already rather fragmented, with distinct traditions in each subject area. In particular: the centre of gravity of theoretical work has moved from statistics departments to mathematics and computer science ones; each application domain has to some extent evolved independently of all the others; and there has been remarkably little interaction between the learning communities and the mathematical programming / optimization community, who have become basic algorithm providers for the field almost without knowing it. Even the name of the field has become fragmented: depending on your emphasis, it can be called pattern recognition, computational or machine learning, or flexible statistical modelling.

The PASCAL NoE will counteract this fragmentation, restructuring European learning research into a coherent community founded on a common core of rigorous theoretical analysis techniques and efficient scalable algorithmic ideas. By defining this syllabus and promoting it at a pan-European level, it will break down the traditional barriers between the various disciplines involved, greatly accelerate the transfer and exploitation of learning-based techniques, and provide a dynamic community of European scientists and engineers with a thorough theoretical and practical grounding in them.

The PASCAL-RTN network will play a privileged role in this. It will provide an initial core of talented young researchers with in-depth training in these techniques. It will use the resulting experience to refine and perfect the training program and syllabus. And it will allow significant research efforts to be focused on key interdisciplinary topics, further breaking down the barriers between disciplines and improving the potential for flow of ideas.

Interdisciplinarity and European uniformity will be assured by study courses leading to PASCAL accreditation and including purpose designed summer schools and tutorial workshops, and by participation in visitors programs to provide longer term exposure to different ideas and points of view. The experience gained will help the PASCAL-RTN scholars to become more responsive to needs and encourage them to apply the new ideas that they have learned in challenging practical problems. European competitiveness will benefit from a deeper understanding of what is possible and wider access to the central core of know-how, and end users will benefit from the improved products and services that result.

B2.3. Planned Recruitment of Early-Stage and Experienced Researchers

PASCAL-RTN will provide a total of 484 person months (40.3 person years) of PhD-level scholarship for 14–15 early-stage researchers at 14 sites, and 318 person months (26.5 person years) of postdoctoral-level scholarships for about 17–20 experienced-stage researchers at 17 sites, for a total cost of approximately 4.2 MEu over 4 years.

To maintain uniform standards, PASCAL-RTN scholarships will be awarded competitively by the network's Scientific Management Committee from among propositions received from host sites for a specific candidate and topic at that site. The committee will be responsible for achieving a good balance across both subject areas and sites, and for assuring equal opportunities for all equally qualified candidates regardless of sex, race and disabilities, and it affirms its commitment to equal opportunity principles. No priority will be given to candidates from other PASCAL or PASCAL-RTN sites: candidates can come from any region so long as they satisfy the EC's criteria for eligibility for an RTN scholarship at the chosen site (including mobility criteria). To prevent wasted recruitment effort, there will be an initial consultation over suitable topics and sites and an initial screening of applications for administrative eligibility.

It will be up to individual host sites to advertise for suitable candidates. We expect them to advertise in the academic vacancies sections of the press, via subject-oriented mailing lists and newsgroups, in the vacancies section of the PASCAL web site, and on the Commission's central Cordis / FP6 employment opportunities site(s). The best responding candidate(s) will be preselected by the advertising site, again according to strict equal opportunities rules, and submitted to the Scientific Management Committee for approval.

Person Months by Site

The table below lists the person months requested by each site. If this is permissable, we would also like to request a small number of unattributed person months (10% of the total network budget, marked 'HELD IN RESERVE' in the table below). These will be managed by the Scientific Committee and the coordinator for the benefit of the network as a whole. For example, they could be used to fund focused research on especially promising opportunities or unexpected problems that arise during the course of the project, or to allow a PASCAL partner who is not currently in PASCAL-RTN to join PASCAL-RTN and fund an unusually promising topic. If this is not permitted, these person months can be deleted.

	Early-stage and experienced researchers to be financed by the contract			**Other professional researchers likely to contribute research effort on the network project**	
Network team	**Early-stage researchers** (person months) (a)	**Experienced researchers** (person months) (b)	**Total** (a+b) (c)	(number of individuals) (d)	(person months) (e)
1. RHU London	-	36	36	1	3
2. INRIA Grenoble	36	6	42	2	6
3. NCRG Aston	-	24	24	1	8
4. LIA Avignon	36	9	45	3	24
5. UPC Barcelona	12	12	24	5	15
6. IDA Fraunhofer	-	40	40	2	8
7. U Bayreuth	36	-	36	2	15
8. EURANDOM	-	24	24	5	6
9. TU Graz	36	-	36	3	24
10. CoSCo Helsinki	36	12	48	5	24
11. MU Leoben	36	-	36	2	9
12. JSI Ljubljana	12	20	32	1	16
13. RRG Oxford	36	6	42	2	6
14. UJM Saint-Etienne	-	12	12	4	12
15. I3S Sophia Antipolis	36	-	36	2	18
16. KTH Stockholm	-	33	33	1	8
17. MPI Tübingen	36	12	48	2	10
18. U Aalborg	-	12	12	1	6
19. UC London	-	24	24	2	6
20. U Milan		12	12	3	4
21. UPMC Paris	52	-	52	2	5
22. ETH Zurich	36	-	36	3	24
HELD IN RESERVE	48	24	72	-	-
TOTAL	**484**	**318**	**802**	**54**	**257**

B3. Quality/Capacity of the Network Partnership

B3.1. Collective Expertise of the Network Teams

This section briefly describes the participating sites and their key researchers, who will be responsible for directly supervising PASCAL-RTN scholars. All of the people listed will also be full participants of the associated Network of Excellence PASCAL.

Partner 1. Royal Holloway College, London, U.K.

RHUL will bring to the project significant expertise in both theory and applications of well-founded statistical learning techniques. Specific strengths are in kernel-based learning methods such as Support Vector Machines, on-line learning techniques and Bayesian belief networks. Projects funded by the EU and the UK funding council EPSRC are examining the development of kernel based learning methods as well as delimiting their applicability as a generic learning paradigm. The outputs of these projects will provide valuable contributions to the proposed work of the current project assessing the applicability of recent learning techniques in novel applications areas.

CVs of Key Personnel:

Prof. John Shawe-Taylor was the coordinator of the ESPRIT NeuroCOLT Working Groups, the KerMIT project. He will act as Site Manager for RHUL. He is researching in Statistical Learning Theory with particular concern to well-founded analysis of generalization based on luckiness properties such as the margin. He has successfully applied learning techniques to fraud detection in mobile telecommunications as part of an EU ACTS project ASPeCT, to document analysis in the KerMIT project and machine vision in the LAVA project. He has published over 100 papers, the majority concerned with the application and analysis of learning systems, as well as a book on Support Vector Machines.

Partner 2. LEAR team, INRIA Rhône-Alpes, Grenoble, France

Dr Cordelia Schmid holds a M.S. degree in Computer Science from the University of Karlsruhe and a Doctorate in Computer Science, from the Institut National Polytechnique de Grenoble (INPG). Her doctoral thesis on local greyvalue invariants for image matching and retrieval received the best thesis award from INPG in 1996. Dr Schmid was a post-doctoral research assistant in the Robotics Research Group of Oxford University in 1996–1997. Since 1997 she has held a permanent research position at INRIA Rhone-Alpes. She received the Habilitation degree in 2001 for her thesis entitled "From Image Matching to Learning Visual Models". Her research interests span computer vision and machine learning, and her recent projects include matching images taken under very different conditions, object recognition and image retrieval, as well as video structuring and indexing. Dr Schmid has been an Associate Editor for the IEEE Transactions on Pattern Analysis and Machine Intelligence since 2001, and she has served on the program committees of several major conferences, notably as an area chair for CVPR'00, ECCV'02, and ICCV'03. She currently co-supervises the European projects VIBES (Video Indexing Browsing and Structuring) and LAVA (Learning for Adaptable Visual Assistants).

Dr Bill Triggs is a permanent researcher (Chargé de Recherche) of the CNRS (Centre National de Recherche Scientifique), working in the LEAR team of the Joint Research Unit GRAVIR in INRIA Rhône-Alpes. An ex mathematical physicist, he also worked on robotics at Oxford University and INRIA before turning to computer vision in INRIA's MOVI research group. His past vision work is mainly on the mathematical foundations of vision geometry and scene reconstruction, including multi-image geometric matching relations and self-calibration, but recently he has begun to work on statistical pattern recognition and machine learning for visual object recognition. He managed the EU LTR project CUMULI on photogrammetry and computer vision, and is currently working on human tracking and reconstruction under the EU video processing project VIBES and on learning methods for visual scene understanding in the EU project LAVA. He has served on many program committees, as area chair for the conferences ECCV'00, ECCV'02, CVPR'03, and he will organize the 2005 International Conference on Computer Vision in Cannes, France. He will serve as an FP6 expert evaluator in April 2003.

For further information and references, see `http://www.inrialpes.fr/movi/people/Schmid` and `http://www.inrialpes.fr/movi/people/Triggs`.

Partner 3. Neural Computing Research Group, University of Aston, Aston, U.K.

Manfred Opper obtained a PhD in physics in 1987 in Giessen, Germany. After that he began to work on the application of statistical physics to the theory of neural networks. After postdoctoral visits at the Ecole Normale Superieure in Paris (1989) and at the University of California at Santa Cruz (1990) he received the habilitation degree in theoretical physics in 1991. In 1992 he was awarded the Physics Prize of the German Physical Society for work on the dynamics of learning in neural networks. After working as a research associate at the universities of Giessen and Wuerzburg, Germany, he was awarded a Heisenberg fellowship in 1994 enabling him to stay at the machine learning group of UC Santa Cruz and at the Department of Complex systems of the Weitzmann Institute in Israel. After substituting a professorship for theoretical physics for one semester at the university of Bayreuth, he joined the Neural Computing Research Group at Aston as a reader in 1997. He has over 100 publications, most of them in the application of statistical mechanics methods to problems of machine learning and other complex systems. His research interests are in the cross connections between statistical physics, computational learning theory, information theory and mathematical statistics. He is currently working on the theory and applications of Gaussian processes and Support-vector Machines and tries to develop efficient inference techniques for probabilistic models [1,2].

The Aston Neural Computing Research Group is an established international research group, focusing on both practical and theoretical aspects of artificial neural networks and other statistically oriented machine learning methods. It currently comprises two Professors, one Reader, five Lecturers and 6 postdoctoral fellows. In addition the group has a full time System Administrator and an Administrative Assistant. The diversity of research activities within the group is both stimulating and supportive, facilitating close and fertile interaction between theoreticians and practitioners as well as direct implementation and examination of ideas, derived by theoretical studies, on real-world tasks.

[1] *Tractable approximations for probabilistic models: The adaptive TAP mean field approach*. Manfred Opper and Ole Winther; Phys. Rev. Lett. 86, 3695-3699 (2001).

[2] *TAP Gibbs Free Energy, Belief Propagation and Sparsity*. Lehel Csató, Manfred Opper and Ole Winther; in: Proceedings of NIPS 2001 (2002).

Partner 4. LIA, Université d'Avignon, France

The Avignon Laboratory for Computer Science (LIA) is a unit of the French National Scientific Research Council (CNRS). Its research activity is centered around the use of machine learning techniques for Information Retrieval, Natural and Spoken language processing.

A stable cooperation with the research laboratories of France Telecom R&D in Lannion and Telecom Italy in Turin has produced interesting results on basic problems encountered in the development of telephone and Internet services. Seventeen permanent researchers and more than 25 PhD students works at the laboratory which occupies 1000 m2 on a new building and has state of the art computer facilities for these research problems.

The laboratory is funded by the French Ministry of Education, the CNRS and with contracts from Industry, especially France Telecom R&D. Recently, LIA has been a contractor for two EU projects, namely SMADA and MTM, the fist one closely related to the actual proposal.

The proposed project will be carried out by the team working on person-machine dialogue headed by Renato De Mori who is also director of the laboratory.

Renato De Mori graduated from Politecnico do Torino in Italy and had a chair in Computer Science at the University of Turin (Italy). From 1985 till 1996 he has been professor and director at the School of Computer Science, McGill University, Montreal, Quebec, Canada. In that period he worked on important research projects with companies including IBM Canada where he served as member of the Executive Advisory board of the Toronto Lab. He has been member of many committees in Canada and the US and is now member of the IEEE Speech Technical Committee. He is Fellow of the Institute of Electrical and Electronic Engineers from 1994 for: Contribution to symbolic and quantitative methods of signal interpretation and understanding.

He is editor in chief of the journal 'Speech Communication' and associate editor of 'Pattern Recognition Letters', 'Computer Speech and Language' and 'Computational Intelligence'. He has also been associate editor of 'IEEE Transactions on Pattern Analysis and Machine Intelligence'.

Two recent publications:

Y. Estève, C. Raymond, R. De Mori et D. Janiszek, On the use of linguistic consistency in automatic speech recognition, 2003 Accepted with minor revisions on the : IEEE Transactions on Speech and Audio Processing.

C. Carpineto, R. De Mori, G. Romano, B. Bigi An information theoretic approach to automatic query expansion ACM Transactions on Information Systems, 19(1):1-27,2001.

Partner 5. UPC Barcelona

The UPC team belongs to a large Department including more than 80 permanent faculty members dedicated to research in virtually all fields of Computer Science, Artificial Intelligence and Information Systems. In particular, they are members of a well established group on algorithms and computational complexity with a sound history of two decades of research.

José L. Balcázar and Ricard Gavaldà are the senior researchers of a subgoup which, for the last ten years, has worked on computational learning theory and data mining, using their strong background on algorithms and computational complexity. In the last five years, the group has produced five Ph.D. theses whose area is, fully or partially, computational learning theory. Among them is that of Jorge Castro, the third doctor involved in this proposal.

José L. Balcázar (Ph.D. UPC, 1984) is a full professor at UPC since 1987. Besides a strong research record on computational complexity, his current research is focused on computational learning theory and algorithmics of data mining. His most recent publications deal with abstract characterizations of efficient query learning, efficient randomized algorithms for training support vector machines, and efficient algorithmics for problems related to basket market analysis. In the past he has been Program Committee chair for the IEEE Computational Complexity Conference and member of the Program Committee of the Algorithmic Learning Theory Conference, and is in the Program Committee of this year's edition of the International Conference on Machine Learning.

Ricard Gavaldà (Ph.D. UPC, 1992) is an associate professor at UPC since 1993. His recent research involves characterizations of efficient query learning, connections between learning problems and algebra, and the design of sampling algorithms useful to scale up data mining and learning algorithms. He is the co-chair of this year's edition of the Algorithmic Learning Theory Conference.

Jorge Castro (Ph.D. UPC, 2002) is an associate professor at UPC. The topics of his recent Ph.D. include combinatorial characterizations of query learning and the design of efficient learning algorithms for classes of boolean functions. is on learning classes of boolean functions and combinatorial characterizations of learning. His most recent interest includes algorithmics of support vector machines.

Partner 6. Intelligent Data Analysis (IDA) group, Fraunhofer Institute for Computer Architecture and Software Technology (FhG FIRST), Berlin

Key people:

Klaus-Robert Müller (10%) received the Diplom degree in mathematical physics 1989 and the Ph.D. in theoretical computer science in 1992, both from University of Karlsruhe, Germany. From 1992 to 1994 he worked as a Postdoctoral fellow at GMD FIRST, in Berlin where he started to built up the intelligent data analysis (IDA) group. From 1994 to 1995 he was a European Community STP Research Fellow at University of Tokyo in Prof. Amari's Lab. From 1995 on he is department head of the IDA group at GMD FIRST (since 2001 Fraunhofer FIRST) in Berlin and since 1999 he holds a joint associate Professor position of GMD and University of Potsdam. He has been lecturing at Humboldt University, Technical University Berlin and University of Potsdam. In 1999 he received the annual national prize for pattern recognition (Olympus Prize) awarded by the German pattern recognition society DAGM. He serves in the editorial board of Computational Statistics, IEEE Transactions on Biomedical Engineering and in program and organization committees of various international conferences. His research areas include statistical physics and statistical learning theory for neural networks, support vector machines and ensemble learning techniques. His present interests are expanded to time-series analysis, blind source separation techniques and to statistical denoising methods for the analysis of biomedical data.

Gunnar Rätsch (10%) received the Diplom in computer science from the University of Potsdam (Germany) in 1998, along with the Jacobi prize for the best student of the faculty of natural sciences.

Three years later, he obtained a Ph.D. in natural sciences with his work on Boosting – being done at the Fraunhofer Research Institute for Computer Architecture and Software Technology in Berlin (Germany). He received the Michelson prize awarded by the University of Potsdam (Germany). From 2001 to 2003 he was working as a postdoctoral fellow in the Research School of Information Sciences and Engineering of the Australian National University in Canberra (Australia). Now he is a research scientist jointly appointed by Fraunhofer FIRST in Berlin and Max-Planck Institute for Biological Cybernetics in Tübingen. His scientific interests and research areas include pattern recognition, on-line learning, bio-informatics and drug-design.

References:

Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K., Schölkopf, B., An Introduction to Kernel-Based Learning Algorithms, IEEE Transactions on Neural Networks, 2 (2), 181-201 (2001)

Rätsch, G., Mika, S.,Schölkopf, B., Müller, K.-R., Constructing Boosting Algorithms from SVMs: An application to One-Class Classification, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 9, 1184–1199, (2002)

G. Rätsch, A.J. Smola, and S. Mika. Adapting codes and embeddings for polychotomies. In Advances in Neural information processing systems, volume 15, 2003.

R. Meir and G. Rätsch. An introduction to boosting and leveraging. In S. Mendelson and A. Smola, editors, Advanced Lectures on Machine Learning, LNAI 2600, pages 119-184. Springer, 2003.

Tsuda, K., Kawanabe, M., Rätsch, G., Sonnenburg, S., Müller, K.-R., A Discriminative Fisher Kernels, Neural Computation, 14 (10), 2397–2414 (2002)

Site facilities: The Intelligent Data Analysis department at Fraunhofer FIRST in Berlin is an academic research group involving 13 experienced researchers and 9 PhD students, working on several aspects of machine learning, bio-informatics and bio-medical data analysis. Its facilities include a biomedical experiments room and a cluster of 16 dual processors computers and about 40 additional computers for large scale computing experiments.

Partner 7. University of Bayreuth, Bayreuth, Germany

Prof. Dr. Klaus Schittkowski (10 %) studied Applied Mathematics at the University of Wuerzburg and received his Ph.D. degrees in 1975 (Dr.rer.nat.) and in 1982 (Dr.habil.). Since 1986, he is working at the Department of Mathematics of the University of Bayreuth. His main research interest is in the development of nonlinear programming algorithms, least squares optimization, implementation and empirical comparative tests of optimization codes, discretization of dynamical systems, automatic differentiation, and large scale topology optimization. His published six books and about 90 articles, and implemented more than 10 codes for nonlinear programming, some of them widely used in industry and academia.

PD Dr. Ch. Zillober (20 %) got his Ph.D. in Applied Mathematics at the Technical University of Munich in 1991. After three years at the Institute of Scientific Computing of the University of Heidelberg, he moved to the University of Bayreuth and became a member of the research group of Klaus Schittkowski. In 2002, he finished his *Habilitation* in Applied Mathematics. His research areas are nonlinear programming, structural mechanical optimization, sequential quadratic programming, and interior point methods. He published about 25 articles and developed an SCP code for very large sale nonlinear optimization.

References:

K. Schittkowski, C. Zillober (2003): Nonlinear programming, Encyclopedia of Life Support Systems (EOLSS), UNESCO, Topic: Optimization and Operations Research, 157-177

K. Schittkowski (2002): Numerical Data Fitting in Dynamical Systems - A Practical Introduction with Applications and Software, Kluwer Academic Publishers

K. Schittkowski, C. Zillober (2002): SQP versus SCP methods for nonlinear programming, to appear: Proceedings of the Workshop Optimization and Control with Applications, Kluwer Academic Publishers

C. Zillober (2002): SCPIP - an efficient software tool for the solution of structural optimization problems, Structural and Multidisciplinary Optimization, Vol. 24, No. 5, 362-371

Partner 8. EURANDOM

Scientists in charge of the research and training program:

Professor Henry Wynn held senior positions at Imperial College, London (Reader), City University, London (Professor and Dean of Mathematics) and the University of Warwick (Professor of Industrial Statistics). In the latter job, he was Director of the Risk Initiative and Statistical Consultancy Unit. This Centre grew to be one of the foremost of its kind in the UK. He joined the London School of Economics (LSE) in 2003 bringing much of his current research activities with him. These range from active collaboration in applied statistics projects particularly with industry and several important theoretical projects. Of the latter, he was one of the pioneers in the new area of computational algebraic statistics and also helped to start the subject of dynamical research. He is Scientific Co-Director of EURANDOM with special responsibility for the statistics program. He has over 130 published papers, three books and is an honorary fellow of several international societies. He is an Honorary Fellow of the Institute of Actuaries and holds the Guy Medal in Silver from the Royal Statistical Society.

Professor Dr. A. van der Vaart, is full professor at the Free University of Amsterdam since 1996. Since 2000 he is coordinator of the Research program Statistics and Modelling Information at EURANDOM. He published over 60 papers. He was editor and referee of several Dutch and international Journals on Statistics, and reviewer of book proposals, and reviewer for NWO, AMS, NSA and NSF grants. >From 1987 on he had several positions as research or teaching assistant in Leiden, Amsterdam and Texas A & M University, was appointed postdoctoral fellow at the Mathematical Research Institute in Berkeley, associate professor in Amsterdam (1992-96) and Professeur Premiere Classe, Université Paris XI (1995-96).

Dr. P. Gruenwald, studied Computer Science in Amsterdam (preliminary exams Computer Science 1989 and Mathematics 1991), where he got his Masters degree in 1994. He started his PhD studies at the CWI, Amsterdam, The Netherlands in 1994, where he defended his thesis under the supervision of professor Paul Vitanyi in 1998. From October 1998 until June 2001 he had a postdoctoral position at the Department of Computer Science, Stanford University, USA and at EURANDOM, in the Complex Statistical Models Group. In July 2001 he was appointed in a research position at the CWI, Amsterdam, The Netherlands on statistics, machine learning and reasoning with uncertainty.

All three scientists are experts in computational learning and mainstream statistics, central to the topic the postdocs are going to work on. Their position at EURANDOM is program coordinator / research fellow for one day a week. In this capacity they discuss plans and progress and provide the postdocs the scientific guidance necessary. Furthermore, the institute itself — with about 24 other young researchers (postdocs) working in five research programs — offers a unique research environment in stochastics with a lively interaction between programs and researchers.

For further information on publications, *etc*., see:

`http://www.cs.vu.nl/~aad/research/index-en.html`
`http://db.cwi.nl/personen/publiek/zoek_show.php4?persnr=413`
`http://www2.warwick.ac.uk/fac/sci/statistics/riscu/research/`

Partner 9. Institut fuer Grundlagen der Informationsverarbeitung, T.U. Graz, Graz, Austria

Biosketch of Wolfgang Maass: Professor Maass received the Ph.D. in Mathematics from the University of Munich in 1974, and the Habilitation in 1978. He spent the following years as a Heisenberg-Fellow of the Deutsche Forschungsgemeinschaft at MIT, the University of Chicago, and the University of California at Berkeley. In 1982 he became Associate Professor, and in 1986 Professor of Computer Science at the University of Illinois at Chicago. In 1991 he moved to Graz, Austria, where he now holds the chair for Grundlagen der Informationsverarbeitung at the Technische Universitaet Graz (Graz University of Technology). Since 1992 he is also head of the new Institute for Theoretical Computer Science in Graz. He spent the winter semester 2002/03 as Guest Professor at the new Mind Brain Institute of the EPFL in Lausanne.

Prof. Maass did most of his earlier research in mathematical logic, computational complexity theory, and computational learning theory. In his recent work he has investigated the computational power and learning complexity of mathematical models for biological neural nets, and has tested theoretical predictions in computer simulations of cortical microcircuits.

Selected References:

W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. Neural Computation, 14(11):2531-2560, 2002.

W. Maass, G. Steinbauer, and R. Koholka. Autonomous fast learning in a mobile robot. In G. D. Hager, H. I. Christensen, H. Bunke, and R. Klein, editors, Sensor Based Intelligent Robots. International Workshop, Dagstuhl Castle, Germany, October 15-25, 2000, Selected Revised Papers, volume 2238 of Lecture Notes in Computer Science, pages 345-356. Springer (Berlin), 2002.

P. Auer, H. Burgsteiner, and W. Maass. Reducing communication for distributed learning in neural networks. In José R. Dorronsoro, editor, Proc. of the International Conference on Artificial Neural Networks – ICANN 2002, volume 2415 of Lecture Notes in Computer Science, pages 123-128. Springer, 2002.

T. Natschläger, W. Maass, and A. Zador. Efficient temporal processing with biologically realistic dynamic synapses. Network: Computation in Neural Systems, 12:75-87, 2001.

W. Maass. Neural nets with superlinear VC-dimension. Neural Computation, 6:877-884, 1994

Site Facilities in Graz: The Institut fuer Grundlagen der Informationsverarbeitung (Institute for Theoretical Computer Science) maintains a cluster of PCs, as well as 2 khepera robots. It currently has 5 Phd students and 2 Assistant Professors working in the group of Prof. Maass on problems in machine learning, neural computation and neural learning, with emphasis on learning principles that can be employed for real-time sensory processing and classification.

Partner 10. Complex Systems Computation Group (CoSCo), Department of Computer Science, University of Helsinki

The Department of Computer Science at University of Helsinki carries out high quality research in its core fields, and has several groups primarily involved in modeling and data analysis. The team in question, The Complex Systems Computation Group (CoSCo) is internationally acknowledged to be a top research group in the area of probabilistic and information-theoretic modeling and their applications. The research areas addressed include stochastic modeling and data analysis — especially with Bayesian networks and related probabilistic model families, such as finite mixture models and Bayesian multinets — information theoretical approaches to inference (MDL/MML), theoretical and empirical validation of on-line prediction algorithms, cased-based reasoning (CBR), and stochastic optimization algorithms, such as simulated annealing and genetic algorithms. Research teams work has both a strong basic research component, being at the intersection of computer science, information theory and mathematical statistics, and an applied component where the results are applied in multiple disciplines from social sciences and medicine to industrial engineering. The recent focus areas in the applied research have been personalization for the Internet, diagnostics for ESA satellites, and modeling for location-aware services.

The key research staff involved in the research on information theoretic learning of graphical models (e.g., Bayesian networks) consists of professor Henry Tirri (University of Helsinki and Stanford University), and 3 senior researchers (Buntine, Rissanen, Myllymäki). The foreseen extent of their involvement is expressed in parenthesis.

Professor Tirri's (10%) special interests are stochastic modeling in computer science and statistics, in particular information-theoretic modeling and Bayesian modeling, game theory in computer science, and learning in natural and artificial systems. Dr. Wray Buntine (15%) is internationally known for his theoretical and practical work in decision trees (e.g., the IND software), and Bayesian graphical modeling. Dr. Jorma Rissanen (20%) is best known for his fundamental work on arithmetic coding and the Minimum Description Length (MDL) Principle. For this research, which in effect place statistics as a chapter in information theory, he has received a number of prestigious awards. Dr. Petri Myllymäki's (20%) special interests are in Bayesian and information-theoretic modeling, in particular in Bayesian networks.

The team has excellent possibilities to support the students to conduct the research in question, modern computing facilities and a stimulating work environment.

Two related publications:

P. Kontkanen, W. Buntine, P. Myllymäki, J. Rissanen, H. Tirri, Efficient Computation of Stochastic Complexity. Pp. 181-188 in Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, edited by Christopher M. Bishop and Brendan J. Frey. Society for Artificial Intelligence and Statistics, 2003.

P. Kontkanen, P. Myllymäki, T. Silander, H. Tirri, and P. Grünwald, On Predictive Distributions and Bayesian Networks. Statistics and Computing 10 (2000), 39-54.

Partner 11. Institute for Computer Science, University of Leoben, Leoben, Austria

This institute was founded in February 2003, headed by Professor Peter Auer. Its main research areas are computational intelligence, machine learning, and their applications to technical and cognitive processes.

O.Univ.-Prof. Dipl.-Ing. Dr. Peter Auer (commitment to PASCAL-RTN: 15%): Peter Auer was born 1964 in Vienna, Austria. He received a Ph.D. in Mathematics in 1992 from the Vienna University of Technology. In 1997 he received the habilitation in Probability Theory and Computer Science from the Graz University of Technology.

Peter Auer has done research in Machine Learning, Neural Networks, Probability Theory, and Symbolic Computation. Since 1993 his main research area is Machine Learning and Neural Networks. He was member of the NeuroCOLT and NeuroCOLT 2 Networks of Excellence under FP4 and FP5, and he is member of the cognitive vision EC-IST project LAVA. A particular research interest are on-line learning methods and statistical methods in learning. Peter Auer was editor of a special issue of *Algorithmica* on "Computational Learning Theory", member of various program committees, and reviewer for a number of scientific journals. From 1986 to 1994 he has been consultant of the Austrian Television Company (ORF) for the forecast of presidential and parliamentary elections.

Two publications:

P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *J. of Machine Learning Research*, 3:397–422, 2002. A preliminary version has appeared in *Proc. of the 41th Annual Symposium on Foundations of Computer Science*.

P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. The non-stochastic multi-armed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, 2003. A preliminary version has appeared in *Proceedings of the 36th Annual Symposium on Foundations of Computer Science.*

Partner 12. Department of Intelligent Systems, Jozef Stefan Institute, Ljubljana

J. Stefan Institute (http://www.ijs.si/), founded 1949, is a research organization for pure and applied research in the natural sciences and technology. At present the Institute, totaling about 700, has a research staff of nearly 450: about 200 of them are post graduates temporarily employed while obtaining their degrees, almost 260 have doctorates, and 100 have permanent professorships or temporary teaching assignments at the Universities (Ljubljana and Maribor). In view of its activities and status, the J. Stefan Institute plays the role of a national institute, complementing the role of the universities and bridging the gap between science and applications.

The Department of Intelligent Systems of the J. Stefan Institute (http://www-ai.ijs.si/) consists of 25 researchers; the head of Department is Prof. Ivan Bratko. It was founded in 1979, initially as an Artificial Intelligence group. In the first ten years, the emphasis was on theoretical research that provided a solid background for later application projects. In 1982, the development and implementation of AI tools started and soon resulted in practical applications. By now, over 60 projects were completed in different domains including business and management, technical domains and medicine. The main research topics of Department of Intelligent Systems are: machine learning, data mining, learning on text and the Web, intelligent systems, information systems, medical informatics, inductive logic programming, automated knowledge synthesis, qualitative modeling, logic programming, decision support, genetic algorithms, knowledge based systems, heuristic programming, knowledge acquisition.

Partner 13. Robotics Research Group, Department of Engineering Science, University of Oxford, U.K.

The research will be carried out within the Robotics Research Group (RRG) of the Department of Engineering Science, University of Oxford. The Department of Engineering Science is large by the standards of most UK universities. It currently comprises 55 academic staff, 100 research assistants, more than 200 research students, and more than 600 undergraduate students. It publishes some 300 scientific papers annually and attracts research support from over 140 companies and agencies. It has consistently achieved top UK government ratings for research excellence. It enjoys the position of a premier graduate school, attracting numerous top-quality graduate students each year, many on scholarships such as the Rhodes, Commonwealth etc. The department provides excellent mechanical, electrical and computing technician services, and access to first class library facilities.

The RRG is one of the largest and best known in its field in Europe, with five faculty and around fifty researchers in total. A substantial part of the group's effort has been invested in the recovery of 3D scene information from multiple images and rendering novel views. Emphasis has been placed on the use of projective geometry and geometrical invariants, as well the use of real-time visual processing. More recently the focus is changing to the use of learning for visual category recognition. Other interests include navigation for mobile robots and wearable computing.

List of personnel:

Professor Andrew Zisserman has researched in the Computer Vision area since 1984. He has written/edited eight books and has co-authored over a hundred papers. In particular he has extensive experience in object recognition (and received the IEEE Marr prize in 1993 for work in this area) and multiple view geometry. He has served on program committees as reviewer and area chair for CVPR, ECCV and ICCV, and was program co-chair for ICCV 98 and ICVGIP 2002. He is the co-chair of ICCV 2005 which will take place in Cannes. He has organized numerous workshops on various aspects of Computer Vision.

Dr Andrew Fitzgibbon is a Royal Society Research Fellow. He has researched in Computer Vision for over 10 years, and also has extensive experience in multiple view geometry. He has served on program committees as reviewer for CVPR, ECCV and ICCV. He was co-organizer of the SMILE workshop at ECCV 2000 and the workshop on dynamic scenes at ECCV 2002.

Two most significant related publications:

R. Fergus, P. Perona and A. Zisserman, 'Object Recognition by Unsupervised Scale-Invariant Learning', Computer Vision and Pattern Recognition, 2003.

A. Fitzgibbon and A. Zisserman, 'On Affine Invariant Clustering and Automatic Cast Listing in Movies', European Conference on Computer Vision, 2002.

Partner 14. Université Jean Monnet, Saint-Etienne, France

EURISE (Equipe Universitaire de Recherche en Informatique de Saint-Etienne) is a small research team: 12 University researchers, 6 doctoral students. It is recognized and financed by the French Ministry of Education and Research. Its main research activities are in Machine Learning and Data mining, with a special trend followed by working on sequential and structural data. Researchers in Eurise are specially active in the field of Grammar Induction (also called Grammatical Inference) with a variety of papers published over the last few years in the main conferences (ICML, ECML, COLT, ICGI, SSPR) and journals (Pattern Recognition, Machine Learning Journal, Journal on Machine Learning Research, ...) where results of this type are published. They have also been following for the past 4 years a line of research devoted to the construction of language models based on grammatical inference. They also have experience in industrial co-operation in Speech Recognition, having been associated on such tasks with the CNET (France Telecom). Between the researchers of Eurise will be more particularly active in the project:

Marc SEBBAN (eurise.univ-st-etienne.fr/∼sebban) is interested in Feature and Prototype Selection using SVM, Boosting or Information Theory. He is currently working on the hard question of reducing noise levels in sequential data. Jean-Christophe JANODET and Marc BERNARD are junior scientists at EURISE interested in grammatical inference over noisy data and learning from structural data. Colin DE LA HIGUERA (eurise.univ-st-etienne.fr/∼cdlh) has been involved with Grammar Induction for some

years. He is the current chairman of the International Community on Grammar Induction. He is organizing a workshop and giving a Tutorial at the next ECML/PKDD conference (at Dubrovnik, in September 2003), on Context-free grammar learning. Researchers from the EURISE are currently developing grammar induction tools whose combination is the main goal for the project. The above researchers will be expected to spend 50% of their research activity, over the next 3 years in actions directly related to the project.

Two representative publications:

Thollard, F., Dupont P., de la Higuera C. Probabilistic DFA Inference using Kullback-Leibler divergence and minimality International Conference on Machine learning, 2000.

Sebban, M., Nock, R., Lallich, S. Stopping Criterion for Boosting-based Data Reduction Techniques: from Binary to Multiclass Problems Journal of Machine Learning Research, MIT Press, 2003, 3, 863-885.

Partner 15. CNRS I3S, Sophia Antipolis, France

I3S (Laboratoire d'Informatique, Signaux et Systèmes de Sophia-Antipolis) is a research unit associated with the CNRS (Centre National de la Recherche Scientifique) and the University of Nice at Sophia Antipolis. It groups more than 100 permanent researchers and about 75 Ph.D. students and Post-Docs, organized in five groups: "Algorithms, Combinatorics and Applications", "Software and Hardware Architectures", "Software Science and Techniques", "Images", "Signal, Robotics, Communication, Control and Optimization (SIROCCO)". Within these groups, four teams form common projects with INRIA Sophia Antipolis.

Luc Pronzato is Research Director at CNRS and vice director of I3S since January 2000. He is also the scientific leader of the team "Techniques for Optimization and Modelling" in the group SIROCCO. The team is composed of 4 staff members and 4 PhD students/Post-Docs/long term visitors, with research activities in robust control, dynamical games, optimization, dynamical systems and statistics (especially experimental design). They are involved in several industrial contracts and national or international academic collaborations on the topics aforementioned. In particular, Luc Pronzato participates, or has participated, in different European projects (BRITE-EURAM, MAST, LTR, GROWTH). He is author or co-author of more than 60 journal papers or book chapters, 50 conference papers and 3 books (Masson, 1994; Springer Verlag, 1997; Chapman & Hall/CRC, 2000). His fields of interest cover system identification, experimental design and optimization, with special attention to sequential methods.

References:

R.Gautier and L. Pronzato. "Adaptive control for sequential design", Discussiones Mathematicae, Probability & Statistics, 20 (1):97-114, 2000.

L. Pronzato. "Adaptive optimization and D-optimum experimental design", Annals of Stat, 28(6):1743-1761, 2000.

L. Pronzato. "Optimal and asymptotically optimal decision rules for sequential screening and resource allocation" , IEEE Trans. on Automatic Control, 46(5):687-697, 2001.

L. Pronzato and E. Thierry. "Sequential experimental design and response optimisation". Statistical Methods and Applications, 11(3):277-292, 2003.

Partner 16. Computational Vision and Active Perception Laboratory, KTH Stockholm, Sweden

The CVAP (Computational Vision and Perception Laboratory) performs research in computer vision and robotics since 1982. The group was formed in 1982 and has today 30 researchers. It is associated with the department of numerical analysis and computing science at KTH, Stockholm.

The research is currently mainly funded through two Swedish sources, TFR, the Swedish Research for Engineering Science, and SSF, the Swedish Foundation for Strategic Research and through a set of grants from the European Union. CVAP has since 1993 a long-term so-called frame grant from TFR and is a partner in two consortia sponsored by SSF, CAS: The Center for Autonomous Systems, and VISIT: Visual Information Technology. VISIT is a national research program hosted by Uppsala University, and involves groups from six Swedish universities. Jointly with the Karolinska Institute, CVAP has a project on the analysis of functional brain images.

Up to 1995 CVAP were partners in three ESPRIT Basic Research Actions: Insight II, VAP II and VIVA. After the ending of these Frame Program III projects, CVAP currently participates in the ESPRIT LTR project Improofs and three networks: RETINA, VIRGO and CAMERA. The group is also a primary node in ECVNet, the European Computer Vision Network.

CV for Stefan Carlsson: Stefan Carlsson joined the computer vision group at KTH in 1991. Between 1992 - 1995 he participated in the European Esprit Basic Research Action 6448, VIVA which was a consortium of some of the major computer vision groups in Europe for doing research on invariance in computer vision. In 1997 he started work in collaboration with partners from the VIVA project in the program IMPROOFS, which involves application of computer vision to to problems in forensic science. He is also involved in the national program VISIT, as a national coordinator for the project "view-synthesis" and also in the project "content based search in image databases".

References:

Recognizing and Tracking Human Action, J. Sullivan and S. Carlsson, Proc 7:th European Conference on Computer Vision (ECCV), Copenhagen, Denmark 2002

Action Recognition by Shape Matching to Key Frames. S. Carlsson and J. Sullivan Workshop on Models versus Exemplars in Computer Vision, Kauai, Hawaii, USA December 14th, 2001

Un-calibrated Motion Capture Exploiting Articulated Structure Constraints, D. Liebowitz and S. Carlsson, Proc 8:th ICCV, Vancouver, Canada, July 2001

Partner 17. Empirical Inference department, Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Key people:

Bernhard Schölkopf (involvement 10% of full time). Bernhard Schölkopf received an M.Sc. in mathematics and the Lionel Cooper Memorial Prize from the University of London in 1992, followed in 1994 by the Diplom in physics from the Eberhard-Karls-Universität, Tübingen, with a thesis written at the Max-Planck-Institute for Biological Cybernetics. Three years later, he obtained a doctorate in computer science from the Technical University Berlin. His thesis on Support Vector Learning won the annual dissertation prize of the German Association for Computer Science (GI). In 1998, he won the prize for the best scientific project at the German National Research Center for Computer Science (GMD).In July 2001, he was elected scientific member of the Max Planck Society and director at the MPI for Biological Cybernetics. Bernhard Schölkopf's scientific interests are in the field of machine learning and perception. In particular, he studies support vector and kernel methods for understanding high-dimensional data.

Jason Weston (involvement 10% of full time). Jason Weston received his B.Sc. in computer science from the Royal Holloway University of London and obtained his PhD in 1999 from the same university under the supervision of Prof. Vladimir Vapnik. He is studying the design of kernel-based learning algorithms for pattern recognition, feature selection, and more complex problems. On the application side, one of his main interest lies in high-dimensional problems of bioinformatics and text mining.

References:

Mika, S., G. Rätsch, J. Weston, B. Schölkopf, A.J. Smola and K.-R. Müller Constructing Descriptive and Discriminative Non-linear Features: Rayleigh Coefficients in Kernel Feature Spaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(4), IEEE Computer Society, Washington, DC (2003)

Weston, J., Chapelle O., A. Elisseeff, B. Schölkopf and V. Vapnik Kernel Dependency Estimation. Advances in Neural Information Processing Systems 15. (Eds.) S. Becker, S. Thrun and K. Obermayer, MIT Press, Cambridge, MA, USA (2003)

Schölkopf, B. and A.J. Smola: Learning with Kernels., 644, MIT Press, Cambridge, MA (2002)

Site facilities: The Empirical Inference department of the Max Planck Institute for Biological Cybernetics in Tübingen is an academic research laboratory involving 14 experienced researchers and 8 PhD students, working on several aspects of machine learning and perception. Its facilities include a psychophysics experiments room, a robotics lab (with robot arm), an optical experiments table (for computer vision experiments), and a cluster of 32 dual processors computers for large scale computing experiments.

Partner 18. Aalborg University, Aalborg, Denmark

Department has around 32 scientists of which 4 are working in area directly associated with current application. Steffen Lauritzen has been Professor at Aalborg University since 1981 and is the author or coauthor of several monographs and articles on graphical models, including articles on their application to learning.

Site facilities include excellent modern computing facilities and electronic access to library, *etc*.

Partner 19. University College, London

The site consists of the Intelligent Systems group in the Department of Computer Science at University College London (UCL), the Gatsby Computational Neuroscience Unit and the Department of Statistics.

At the Gatsby Computational Neuroscience Unit, the site member is Dr Zoubin Ghahramani who is a Lecturer working on statistical approaches to machine learning. He has published over 60 peer-reviewed papers in a variety of topics relating to machine learning, Bayesian statistics, and sensorimotor control, and has served on the editorial board of Journal of Machine Learning Research and on the organizing and program committees of several international conferences (NIPS, UAI, AIStats, ICML). He leads the machine learning group at the Gatsby Unit which comprises a team of researchers studying graphical models, Bayesian statistics, unsupervised learning, kernel methods, Markov decision problems, and other core topics in learning theory. The Gatsby Unit is closely allied with the Machine Learning group in Computer Science, and the two together provide an ideal training site with multiple seminar series, graduate courses, and journal clubs.

At the UCL Computer Science Department, the following lecturers are site members: Dr. Mark Herbster, Dr. Massimiliano Pontil, Dr. Joel Ratsaby.

Dr Mark Herbster joined the Computer Science Department in May 2001where he is currently a lecturer and the director of the MSC program in intelligent systems. His main research interest is in machine learning in particular the on-line learning framework. He has published in the Journal of Machine Learning, the Journal of Machine Learning Research, and has recently been a member of the COLT program committee. Recently, he has been studying the computational complexity and the quality of approximation by on-line methods using a variety of orthogonal systems (reproducing kernels and wavelets). Currently, he is supervising a doctoral student in the development an of on-line basis selection technique with multidimensional wavelet packets.

Dr. Pontil joined the Computer Science Dept in January 2003. His main research interests involve machine learning and pattern recognition. He has published extensively in these areas. In particular, over the past two years Massimiliano research studies have involved different aspects of learning theory and its applications. In particular: with Prof. Charles Micchelli (SUNY, Albany, USA) he is working on approximation problems in learning theory; with Prof. Theodoros Evgeniou (INSEAD, Fontainebleau, France) and Dr. Andre Elisseeff (Max Planck Institute, Tübingen, Germany) he developed a new framework for studying randomized learning algorithms such as bagging and boosting; with Dr. Bernd Heisele (Honda Research Labs, Cambridge, USA) and Prof. Tomaso Poggio he proposed a method for learning a hierarchy of kernel classifiers and applied the method to object recognition and detection tasks in computer vision; with Prof. Paolo Frasconi (University of Florence, Italy) he studied model selection methods for multiclass classification systems using kernel machines techniques and applied these methods to fingerprint classification problems; during a long visit at City University of Hong Kong, under invitation of Prof. Steven Smale (now at UC Berkeley, USA), he investigated some aspects on the foundation of learning theory.

Dr. Joel Ratsaby joined the group as a lecturer in the Department of Computer Science in October 2002. He has published over 10 peer-reviewed papers in computational learning theory, pattern recognition, functional approximation theory, statistical model selection and stochastic on-line learning. He obtained a Ph.D. from the Moore School of Electrical Engineering at the University of Pennsylvania, did a two-year postdoctoral in computational learning theory at the Technion, Israel Institute of Technology, and later in the industry, he worked on advanced data mining applications in various commercial projects, was the AI group leader at a start-up company which developed an AI-based recommendation engine for e-commerce. Since joining UCL in October 2002, Dr. Ratsaby has been teaching the advanced MSc course in Intelligent Systems titled “Supervised Learning” and has been doing research on

learning multi-category pattern classification.

Partner 20. University of Milan

Nicolò Cesa-Bianchi received his Ph.D. degree in informatics from the University of Milan (Italy) in 1993. Since 1998, he is associate professor in computer science at the University of Milan. His main research interests are in the areas of machine learning, statistical pattern recognition and on-line algorithms. On these topics, he has published several papers on international scientific journals, including Journal of the Association for Computing Machinery, Annals of Statistics, Machine Learning Journal, SIAM Journal on Computing, Journal of Computer and System Sciences, IEEE Transactions on Neural Networks, and many others. He has been serving several times on the program committee for the conference on Computational Learning Theory and for the conference on Algorithmic Learning Theory. He has been program chair of the thirteenth annual conference on Computational Learning Theory (COLT 2000) and of the 13th International Conference on Algorithmic Learning Theory (ALT 2002). Since 2001, he is member of the steering committee for the Conference on Learning Theory.

References:

N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms Journal of Computer and System Sciences, 64(1):48-75, 2002.

N. Cesa-Bianchi, Y. Freund, and R.E. Schapire. The nonstochastic multiarmed bandit problem SIAM Journal on Computing, 32(1):48-77, 2002.

N. Cesa-Bianchi and G. Lugosi. Potential-based algorithms in on-line prediction and game theory Machine Learning, 51(3):239-261, 2003.

Partner 21. Université Pierre et Marie Curie, Paris

Patrick Gallinari is Professor at University Paris 6, Computer science Dept, and vice director of the Dept. His research interests are statistical machine learning, textual information retrieval, user modeling, and handwriting recognition. LIP6 was the French node of Neuronet (Network of Excellence on Neural Networks), and P.G. was in charge of the research committee of this network for 4 years. He is also coordinator of a French excellence network on machine learning.

Florence d'Alché-Buc (10%) received a PhD in Computer Science from University of Paris XI (Orsay) in 1993 and diploma of 'Habilitation à diriger des recherches' from University of Paris VI in 2001. She was researcher at Philips France research for 2 years and from 1995 has been an assistant professor at university of Paris VI. Her research interests cover constructive algorithms, ensemble methods, generative and graphical models and kernel-based algorithms. Recently she moved from text-mining applications to gene expression modeling.She is author or co-author of 30 publications in journals or conference proceedings and of two patents. She acts as referee for a number of international conference (ICANN, NIPS) and international journals (IEEE).

Partner 22. Eidgenoessische Technische Hochschule, Zurich, Switzerland

Luc Van Gool studied Electrical Engineering at the University of Leuven, where he also obtained his PhD degree with research into the use of invariance theory for computer vision. He is a professor at the University of Leuven, Belgium, since 1992 and a full professor at the ETH in Zurich. Switzerland, since 1998. At both places he leads the computer vision research activities. He also is an associated professor at the University of Purdue in the US. He is a co-founder and director of the company Eyetronics NV, which is active in the field of 3D acquisition and animation and which has been created early '98. He is a member of the editorial boards of several journals, including the International Journal on Computer Vision and Pattern Analysis and Applications, He received several prizes and awards, including a David Marr Prize, a Henri Ford European Conservation Award, a Golden Eye Award, and two TechArt Prizes. His main research interests include object recognition, shape description and reconstruction, texture analysis and synthesis, and scene understanding. He also has a particular interest in the interaction of computer vision with other fields such as psychophysics, neurophyisiology, and archaeology. He is involved in several, major Swiss research projects.

He currently is also involved in 9 European projects. An overview of publications can be found at `http://www.vision.ee.ethz.ch/db_queries/publication.en.html`

B3.2. Intensity and Quality of Networking

PASCAL-RTN is designed to ensure very thorough networking, broad interdisciplinarity and frequent collaboration between sites. As explained in section B2.1, all participants of PASCAL-RTN will also be considered to be nominal members of the Network of Excellence PASCAL, and will thus receive the full benefit of its summer schools, workshops, and other activities. All PASCAL-RTN scholars will be required to play an active role in PASCAL, and their participation and study program will be monitored by specially appointed mentors from other sites. PASCAL-RTN PhD scholars will be expected to participate sufficiently to gain the PASCAL PhD accreditation, and they will also have to spend time at other PASCAL sites as part of their overall training (a minimum of 4 months over the 3 year studentship, and preferably 6-8 months).

B3.3. Relevance of Partnership Composition

The partners of PASCAL and PASCAL-RTN have been carefully selected for their complementarity and their potential to contribute to our overall goal of reshaping the scattered research on computational learning and statistical pattern recognition in Europe into a single coherent discipline centring on well-founded methods and algorithms. The networks have the explicit goal of breaking down interdisciplinary barriers and cementing strong research collaborations between computational learning, statistics, optimization theory and the individual application domains (vision, text/web, speech, user interfaces).

Several interdisciplinary connections already exist among the PASCAL partners, in many cases built up during previous collaborative EU research projects. For example, the coherence, transparency and international status of the European computational learning community was greatly enhanced by the successful NeuroCOLT and NeuroCOLT II networks, in which many of the computational learning partners of PASCAL participated. Similarly, several partners of PASCAL are associated with EURANDOM, the European research institute on random processes.

Several PASCAL partners are also collaborating on application-oriented projects, most notably the projects KerMIT on kernel methods for text analysis and LAVA on learning methods for visual scene understanding.

All of the above projects involve both academic and industrial partners, and one of the goals of PASCAL/PASCAL-RTN is to significantly enhance scientific-industrial collaboration in the PASCAL subject areas.

All participants of PASCAL-RTN belong to EU members (85%) or associated states (15%). No involvement of partners from non-associated states is foreseen.

B4. Management and Feasibility

B4.1. Proposed Management and Organizational Structure

To ensure good inter-network coordination, PASCAL-RTN will be managed by the same scientific and executive committees as the Network of Excellence PASCAL. The structure is as follows.

- There will a core **Scientific Committee** responsible for managing and coordinating the scientific program, attributing scholarships, and verifying that training targets and research milestones are met. This committee will comprise a chair (the network coordinator, Prof. John Shawe-Taylor of Royal Holloway College, London), the managers of the scientific programs, who will also act as representatives for core sites, and an Executive Officer (to be appointed).

- Day to day management will be assured by a smaller **Executive Committee** based at the coordinating site. This will be chaired by the Executive Officer, and it will include representatives of the coordinating institution (Prof. Frances Robinson of RHUL) and appropriate administrative staff (Accounts: Jane Read).

- The overall strategic direction of the networks will be set by a **Strategic Advisory Council** comprised of senior scientific figures from outside the network, representatives of national and international funding agencies, and industry representatives. The Council will also monitor ethical and equity issues relating to the network, and serve as an appeals panel in the case of internal disputes that can not be resolved by the Scientific Committee.

The Executive Committee will manage any unpaid network funds, monitor overall budgets and expenditure on the various programs, and be responsible for reporting, quality and risk management, and overall audit systems. Scholarship funding will be paid to host sites within one month after the Commission has accepted the relevant reports and cost claims relating to their scholarships and transferred the corresponding funds to the Coordinator. Advances will be paid according to the Commission's rules, following completion of the scholarship contract application and its acceptance by the Coordinator and the Commission, and transfer of the associated funds from the Commission to the Coordinator.

Management related budget received by the network will cover part of the salaries of the Executive Officer (we estimate 15% of full time) and the other administrative staff involved at the coordinating site, and also any expenses relating to external audits required by the Commission. .

Each host site will be responsible for maintaining auditable contractual and accounting information (including detailed time sheets) relating to the scholars employed and the expenses claimed.

Dissemination Program

The results of the PASCAL-RTN project will be disseminated via PASCAL's coordinated dissemination program — which includes research and tutorial workshops, summer schools, scientific and industrial visitor programs, and virtual science park and brokerage of expertise programs designed to promote technology transfer to industry — and also through the usual scientific channels and via each sites industrial contacts.

Intellectual property ownership will be regulated by the contract and by a shared PASCAL/ PASCAL-RTN Consortium Agreement, which will be based on the standard model agreement supplied by the Commission when this becomes available.

B4.2. Management Know-how and Experience of Network Coordinator

The overall coordinator of PASCAL and PASCAL-RTN will Prof. John Shawe-Taylor of Royal Holloway College, London (RHUL). RHUL will also undertake the financial management of PASCAL-RTN. However to distribute the workload, Dr Bill Triggs of the CNRS representing INRIA Grenoble will act as day-to-day scientific coordinator of PASCAL-RTN.

RHUL has a substantial amount of experience in coordinating previous national and European research contracts and networks, and although a suitable Executive Officer still needs to be appointed, the management and administrative structures needed for PASCAL and PASCAL-RTN are already largely in

place. In particular, Prof. Shawe-Taylor and RHUL coordinated the successful NeuroCOLT and NeuroCOLT II Networks of Excellence under FP4 and FP5, which involved a subset of the partners of PASCAL.

Dr Triggs coordinated the FP5 project CUMULI and serves as a site co-coordinator and workpackage manager in the projects VIBES and LAVA.

B4.3. Management Know-how and Experience of Network Teams

Partner 1, RHU London. As co-coordinator of the ESPRIT NeuroCOLT Working Groups in Neural and Computational Learning (see home page http://www.neurocolt.org), Royal Holloway has established itself as one of the lead European sites in the area of Computational and Statistical Learning. The European Conference in Computational Learning Theory (EuroCOLT) was initiated jointly by Royal Holloway and LSE in 1993 and has since developed into a well-recognized international conference, recently merging with the main COLT conference that will now alternate its venue between Europe and the US. More recently Royal Holloway has become the coordinator of the KerMIT project (see http://www.euro-kermit.org) that is taking some of the theoretical results of the NeuroCOLT project into the analysis of text and image documents in collaboration with Xerox and Reuters. Hence, it has considerable experience in the management of European research projects in the area of well-founded computational learning. This includes long-standing relationships with many of the partners in the current proposal.

Partner 2, INRIA Grenoble. The members of the INRIA team have collaborated on a number of successful EU computer vision projects, including SECOND, CUMULI, IMPACT, VIBES and LAVA. Bill Triggs coordinated the project CUMULI and will serve as scientific coordinator of the current PASCAL-RTN application and as co-manager of the PASCAL NoE's Thematic Programs action.

Partner 3, NCRG Aston. The Neural Computing Research Group at Aston University had the following grants within EC networks: (*i*) STIPCO European network on: 'Statistical physics of Information Processing and Combinatorial Optimization' (EC Research Training Network program, grant HPRN-CT-2002-00319). (*ii*) BLUEWATER: Computerized video camera image analysis for monitoring pollution in water (grant IST-1999-10388). (*iii*) NEUROSAT: Neural Networks for remote sensing data (FP4). (*iv*) NAOC: Neural Algorithms for Ocean Colour (FP5).

Partner 4, LIA Avignon. In the last years, the laboratory has successfully managed more than ten contracts per year including two EU projects, SMADA and MTM.

Partner 5, UPC Barcelona. The UPC team has been an active site in several EU-funded projects, including:
– the NeuroCOLT I and II Working Groups on Computational Learning theory;
– (as part of the larger algorithms and complexity group) a series of four ESPRIT projects ALCOM, ALCOM II, ALCOM-IT and ALCOM-FT, which have run without interruption from 1988 until now;
– several bilateral projects with French and German Universities, in the form of Integrated Actions co-funded by the Spanish and the French or German Research Ministries.

In addition, the team has participated (and, mostly, directed) numerous projects funded by Spanish research agencies, at least three of which had computational learning theory as its main topic.

Partner 6, IDA Fraunhofer. The Intelligent Data Analysis (IDA) research group of the Fraunhofer Institute for Computer Architecture and Software Technology in Berlin (FhG FIRST) is concerned with learning systems for data analysis. In particular, IDA is developing tools for high-dimensional multivariate statistics based on methods originally developed in the field of classical statistics and, more recently, in the neural networks, signal processing and machine learning community. The major objective of IDA is to pursue research all the way from theory to application. To this end, IDA is contributing to the theoretical development of new data analysis techniques, to their implementation according to software-engineering standards, and, finally, to their industrial or academic application.

IDA is developing algorithms and software tools for (a) classification and regression with neural networks and support vector machines (e.g. within the EU NeuroCOLT and NeuroCOLT II projects), (b) time series analysis (within EU STORM Project) and (c) blind source separation (EU BLISS project). Algorithms of the IDA group are used in various projects in the application fields: (1) hand-written character recognition (e.g. within an project with AT&T), (2) classification of charm-quark events from

high-energy physics experiments at CERN, (3) analysis of financial data, (4) bio/neuro-informatics, i.e. the analysis of (i) neurological signals like EEG (Electro-Encephalogram), MEG/MNG (Magneto-Encephalogram, Magneto-Neurogram), (ii) protein structures or (iii) TIS & Splice Site recognition in DNA, and (5) acoustic source separation. The group has successfully mastered several BMBF, DFG and industrial projects and is therefore very experienced in managing research and application projects.

Partner 7, U Bayreuth. Besides research projects and seminars funded by national research organizations like DFG, BMBF, MINERVA and by NATO ASI grants, the research group was involved in the EC program COMMETT from 1988 to 1993 and is now node member of the MACSI Network of Excellence.

Partner 8, EURANDOM. EURANDOM is very young institute that only started operating in 1998. As an institute it has not much experience with managing EU projects. By now in the international context we have some experience with Marie Curie Fellowships and with ESF grants. Individual senior researchers at EURANDOM have either a lot of experience with applying and managing grants on the national level (NWO) or have a lot of experience with EU grants (Thematic Networks, Large conferences, *etc*.). EURANDOM is one of the core sites in the PASCAL Network of Excellence where EURANDOM will manage the Workshop Program, the "Computational, information-theoretic learning with statistics" Program (shared activity), the Conference organization Program and the thematic program "Linking learning and statistics with optimization" (shared activity).

Partner 9, TU Graz. Prof. Maass has 3 years of experience in the management of the site in Graz for the EU-Project COLT I, and 4 years of management experience for the EU-Project COLT II.

In addition he has experience in project management for several research grant from the the National Science Foundation of the USA, the Austrian Science Fund, the Austrian National Bank, and the Austrian Government. He has been the head of the Institute for Theoretical Computer Science for 11 years.

Partner 10, CoSCo Helsinki. The CoSCo group has participated in two research networks within Europe: Neural and Computational Learning (NeuroCOLT) and Highly Structured Stochastic Systems (HSSS). In addition to this, the group was also involved in the EU project "Probabilistic Modeling of Baltic Salmon Stocks" (PROMOS) coordinated by the Finnish Game and Fisheries Research Institute, Helsinki, and in the ESA funded project "Software for Autonomous Satellites (SFAS)", coordinated in Finland by Space Systems Finland. In addition the Director of the group, Henry Tirri has 10 years experience in European union projects. CoSCo is also currently preparing as a coordinator a related STREP-project for the first call of FP6.

Partner 11, MU Leoben. Peter Auer was a member of the NeuroCOLT and NeuroCOLT 2 Networks of Excellence under FP4 and FP5, and he is site coordinator of the cognitive vision EC-IST project LAVA.

Partner 12, JSI Ljubljana. The research work has been performed within projects funded within European research projects, as well as by the Ministry of Research and Technology of Slovenia. The following project were funded as European research projects: KDNet (2002-2004), KMForum (2001-2003), Sol-Eu-Net IST-1999-11495 (2000-2002), INCO Copernicus - Concede (1998-2000), INCO-Copernicus Information System for Employment Optimization on Internet (1997-99), COST-258 The Naturalness of Synthetic Speech (1997-99), ESPRIT IV Network of Excellence in Evolutionary Computation (EvoNet) (1996-98), TELRI, Trans-European Language Infrastructure (1995-98), ESPRIT III and IV Network of Excellence in Computational Logic (1991-99) ESPRIT IV LTR Inductive Logic Programming II (1996-98), MULTEXT-EAST, Multilingual Texts and Corpora for Eastern and Central European Languages (1995-97), ESPRIT III and IV European Network of Excellence in Machine Learning (1992-97), ESPRIT III Basic Research Project Inductive Logic Programming (1992-95), PECO92 Inductive Logic Programming Pan-European Scientific Network (ILPNET) (1993-96), a Copernicus Network of Excellence in Inductive Logic Programming (ILPNet2) (1998-2001), COST-233 Prosody in Synthetic Speech (1989-95), ESPRIT II project ECOLES Development of Representation for Machine Learning from Imperfect Information (1989-92), COST 13 project AI and Pattern Recognition (1985-87).

Partner 13, RRG Oxford. The group has been involved with a number of previous EC projects including Esprit BRA's (VIVA,IMPACT,IMPROOFS) as well as ACTS Project AC074 VANGUARD. All of these projects completed successfully. In addition for both IMPROOFS and VANGUARD patents

and licensed software were generated. The group is currently involved in the FET projects VIBES and CogViSys.

Partner 14, UJM Saint-Etienne. The EURISE team has not been involved in previous EU project experiences: the team is fairly recent, having been created (from nothing) 5 years ago. The members of the team have nevertheless experience in international research actions such as Sokrates/Ersamus (for student interchange), joint workshop or conference organization, or scientific animation. For instance, Colin de la Higuera is the current chairman of the International Community on Grammar Induction, supervising thus edition, workshop organization and other promotional actions for the field.

Partner 15, I3S Sophia Antipolis. Luc Pronzato has participated in the following projects:
1) joint TEMPUS program (JEP 2538) System identification and adaptive control, 1991-1994,
2) HCM project SIMONET (System identification and modelling network),
3) MAST project MAUVE (Miniaturized and Reconfigurable Instrumentation for Multipurpose Survey with a Mini AUV), 1996-1998,
4) LTR project NARVAL (Navigation of Autonomous Robots via Active Environmental Perception), 1996-1998,
5) BRITE-EURAM project CE2 (Computer Experiments for Concurrent Engineering), 1996-1999 (L. Pronzato was the person in charge of the project for I3S).

He is currently participating in the following two projects (and is the person in charge for I3S):
6) GROWTH project ONE (Optimization Methodologies for Networked Enterprises), 2002-2004,
7) GROWTH project TITOSIM (Time to Market Reduction via Statistical Information Management), 2001-2004.

Partner 16, KTH Stockholm. CVAP has been involved in many previous EU research projects. Since Dec 2000, Stefan Carlsson is coordinating the IST-FET project VIBES, Video Browsing Exploration and Structuring, involving 6 partners from different countries.

Partner 17, MPI Tübingen. The department in Tübingen is new so there is little previous experience of project management, but the senior scientists have both participated in previous EU projects, notably the NeuroCOLT networks.

Partner 18, U Aalborg. Participant in EU projects DRUMS and BaKE. Chairman of scientific 4 year program (Highly Structured Stochastic Systems) under the European Science Foundation.

Partner 19, UC London. The Gatsby Unit and the Computer Science department at UCL have participated in both EU and other international projects. The UCL site manager, Dr Ghahramani, has participated in the EU Marie Curie Training Site for the Institute of Movement Neuroscience, in a UK EPSRC network for speech and sound processing, and in a large US DARPA project on intelligent agents. Dr Ghahramani has also managed grants for the US National Institute of Health, manages a team of researchers in machine learning funded by the Gatsby Charitable Foundation, and has organized conferences, workshops, and summer schools.

Partner 20, U Milan. Nicolò Cesa-Bianchi has been team manager and work area manager for the ESPRIT Working Group EP 27150 on Neural and Computational Learning II (NeuroCOLT II, 1998-2001). He is currently site manager for the EU project on Kernel Methods for Text and Images (KerMIT) IST-2001-25431. He is core site manager and manager of the the pump priming program in the PASCAL NoE proposal.

Partner 21, UPMC Paris. The Machine Learning team at Université Pierre et Marie Curie is concerned with statistical machine learning and focus on a few main application areas such textual information access and information retrieval, user modelling. The team has been involved in a number of European and French research projects.

Partner 22, ETH Zurich. ETH has been a partner in several European projects, *e.g*. Cimwos, CogViSys, MESH, and 3D-Murale under the 5FP. Luc Van Gool who will supervise the work planned at ETH has coordinated several European projects, including VIVA, Vanguard, Impact, and Improofs.

B5. Relevance to the Objectives of the Activity

The thematic coverage and critical mass needed for PASCAL's central goal of building an integrated European research community focused on well-founded pattern analysis, statistical modelling and computational learning and their applications can only be achieved by a concerted effort at the European level. Although they share common theoretical roots, these techniques cover a wide range of different methods and applications and the necessary thematic scope and depth are not available at a national level in any European country. PASCAL's activities are designed to permanently reduce thematic and national fragmentation and promote interdisciplinarity and technology transfer in these areas. PASCAL-RTN is an important component of this strategy. It will provide a core of young European researchers with a rigorous, thoroughly interdisciplinary training and with in-depth working experience in some of the best research teams in Europe in this area. It will also break down thematic and national barriers by providing an opportunity for focused interdisciplinary research on key 'watershed' topics. PASCAL's curriculum reorganization program (which will be tested first on PASCAL-RTN scholars) will help to consolidate its community-building activities, providing pan-European standards and permanently reshaping the way that these subjects are studied in Europe. And PASCAL's coherent set of training and transfer activities will improve the uptake of these ideas in industry and strengthen European competitiveness in these key strategic areas.

B6. Added Value to the Community

Less-favoured Regions, Candidate Countries and Associated States. PASCAL-RTN involves one partner from an Eastern European candidate country — Department of Intelligent Systems, Jozef Stefan Institute, Ljubljana, Slovenia — and one Swiss partner — ETH (Eidgenoessische Technische Hochschule), Zurich. Besides these, the sister Network of Excellence PASCAL, which will provide the framework within which PASCAL-RTN collaboration and training takes place, also involves several partners from the Associated State Israel: the Hebrew University of Jerusalem; the Technion, Haifa; Tel Aviv University and Bar Ilan University.

Gender Issues. PASCAL and PASCAL-RTN are committed to providing equal opportunities for all equally qualified candidates, regardless of gender, race or disabilities. Women are still noticeably underrepresented in the networks' subject areas, particularly in their more theoretical core subjects. One of the main goals of the sister network PASCAL is to produce a lasting, pan-European reorganization of these subject areas, and we hope that the new courses and training materials developed (for which PASCAL-RTN scholars will serve as the initial guinea pigs!) will help to attract highly-qualified women into these fascinating research areas. Perhaps even more significantly, the large-scale deployment of the highly adaptive learning based technologies developed in PASCAL/ PASCAL-RTN will help to make many user interfaces, sensing systems and data mining methods more flexible and more responsive to different needs and interaction styles. This will help to narrow the 'information' and 'technology' gaps and reduce the effects of subtle design biases that favour the over-represented sectors of the community (here, presumably educated adult male technologists).

Attractiveness and competitiveness of Europe. Machine learning and pattern recognition are quickly becoming economically and strategically important areas owing to the explosive growth of the information economy and the impressive performance gains of recent machine learning methods, and this trend will accelerate in the foreseeable future. Europe currently leads the world in many areas of learning and pattern recognition theory, but North American firms still tend to dominate in the commercial applications of these techniques. PASCAL will help to change this by permanently restructuring the European research and development communities in these areas, providing a unified basic curriculum, easier access to advanced training, improved interdisciplinarity and enhanced transfer between theoreticians, applications people and end users. PASCAL-RTN has two specific roles in this: firstly, to provide a core of talented young European researchers with well-targeted, in-depth, interdisciplinary training; and secondly, to break down interdisciplinary barriers by focusing research energy on a number of carefully selected 'watershed' areas that currently separate disciplines and hamper the flow of ideas and techniques between theory and applications.

Coordination with national and international activities. Many European countries have national societies or interest groups dedicated to pattern recognition and machine learning. Most involve both academic and industry people. Many of PASCAL's and PASCAL-RTN's partners participate actively in their national groups, and such groups will be an important vehicle for diffusion of PASCAL know-how into industry. PASCAL's activities will help to restructure these groups and make them more effective, and its work on defining and reorganizing the basic curriculum in these areas will involve close collaboration with higher education institutes across Europe.

B7. Indicative Financial Information

Each partner intends to claim all eligible appointment expenses for the scholars employed (salaries at the official nationally-adjusted EC rates; employer social security costs; mobility, travel, training and career allowances; *etc*.), and overheads at the maximum allowed rate of 10% of direct expenses.

Category A expenses: This covers expenses incurred by researchers who are not directly employed by the RTN, but who need to supervise or collaborate with RTN scholars to achieve the goals of the project. In the case of PASCAL-RTN, there will be no costs in this category. The sister Network of Excellence PASCAL will be the main vehicle for specialist training and collaboration in PASCAL-RTN, and all PASCAL-RTN scholars will be expected to participate actively in PASCAL's programs. The scholar's costs for this will be covered from the standard training and travel allowances attached to their scholarships, and as their supervisors are all PASCAL participants, the supervisors' costs in this category can be met from their PASCAL budgets. This organization (keeping PASCAL-RTN strictly for scholarship-related expenses, not activity-organizing ones) will help to ensure that all PASCAL-related activities remain as widely available as possible: many PASCAL participants will not participate in PASCAL-RTN because they do not need EU support for associated PhD's or postdocs, but they will still receive full access to the associated training programs if these are organized by PASCAL rather than PASCAL-RTN.

Category B expenses: This covers expenses related to the scientific administration of the RTN. Again there will be no claims in this category in PASCAL-RTN. PASCAL-RTN management committee meetings will coincide with PASCAL ones, and category B expenses relating to these will usually be claimed under PASCAL. Other management costs of PASCAL-RTN are likely to be minor, and will be covered under the standard 10% overhead that will be claimed by each partner.

Category C expenses: We were not able to calculate this column analytically as it is not yet clear what level of external audit the Commission will require during FP6, and therefore how much the needed administration and auditing will cost. For now, we have claimed the maximum allowable rate (7% of the total Community contribution) for all sites. However we expect that the final analytical management costs will be substantially less than this and largely concentrated at the coordinating site, so it is likely that we will be able to significantly reduce these figures at contract negotiation time.

Category D expenses: This category covers other eligible expenses including capital equipment. To keep the total budget down we asked all partners to keep this to a strict minimum. However up-to-date computing equipment is a prerequisite in this research area because pattern recognition and machine learning are computationally very demanding, especially when large-scale real-world problems are considered. Mobility, research networking and transfer activities such as demos also presuppose up-to-date portable computers. For this reason, we have allowed partners to claim up to 1,000 € per 12 months of awarded scholarship (payable at the start of the scholarship) towards the provision of user-level computer equipment for the scholars. All of the entries in column D are for such equipment.

Indicative Financial Information on the Network Project (excluding expenses related to the recruitment of early-stage and experienced researchers)				
Network Team	Contribution to the research / training / transfer of knowledge expenses (€)		Management activities (including audit certification) (€)	Other types of expenses / specific conditions (€)
	(A)	(B)	(C)	(D)
1. RHU London	-	-	7%	3000
2. INRIA Grenoble	-	-	7%	3000
3. NCRG Aston	-	-	7%	-
4. LIA Avignon	-	-	7%	-
5. UPC Barcelona	-	-	7%	-
6. IDA Fraunhofer	-	-	7%	3000
7. U Bayreuth	-	-	7%	3000
8. EURANDOM	-	-	7%	-
9. TU Graz	-	-	7%	3000
10. CoSCo Helsinki	-	-	7%	3000
11. MU Leoben	-	-	7%	3000
12. JSI Ljubljana	-	-	7%	3000
13. RRG Oxford	-	-	7%	3000
14. UJM Saint-Etienne	-	-	7%	-
15. I3S Sophia Antipolis	-	-	7%	-
16. KTH Stockholm	-	-	7%	3000
17. MPI Tübingen	-	-	7%	4000
18. U Aalborg	-	-	7%	-
19. UC London	-	-	7%	-
20. U Milan	-	-	7%	-
21. UPMC Paris	-	-	7%	3000
22. ETH Zurich	-	-	7%	3000
Totals	**-**	**-**	**7% of budget**	**40 000 €**

B8. Previous Proposals and Contracts

The current PASCAL-RTN proposal is closely associated with, and complementary to, our FP6 Network of Excellence proposal PASCAL, which will be submitted to the "2.3.1.6 Multimodal Interfaces" objective of call FP6-2002-IST-1 (closing 24 April 2003). We do not yet have a submission-ready version of the PASCAL workplan, but a complete draft is available on demand.

The PASCAL networks are in some sense a continuation (with the same coordinator, but with greatly extended subject area and scope) of the successful FP4 and FP5 networks on Computational Learning Theory for neural networks and kernel based learning, NeuroCOLT I and II.

Several partners of PASCAL and PASCAL-RTN are currently collaborating on targeted RTD projects under FP5 and FP6. The most notable examples are the projects KerMIT and LAVA, respectively on kernel based learning for text analysis and for visual object recognition. As LAVA is a Cognitive Vision project, its partners also nominally belong to the FP5 NoE ECVision, although as far as we know none of them have ever actually requested funding from this source.

Besides the above, we are aware of the following existing or proposed involvements of PASCAL-RTN participants in FP5 and FP6 activities in the PASCAL-RTN subject areas. We do not consider any of these to be excessive given the thematic differences and the fact that the main goal of PASCAL and PASCAL-RTN is training and community-building, not application-directed targeted research:

Partner 1, RHU London. A related Marie Curie fellowship was awarded to Juho Rousu starting 2003 working on relationships between Bayesian Networks and kernel methods. There will be no explicit overlap with the subject of the current work.

Partner 2, INRIA Grenoble. The Joint Research Unit to which we belong, GRAVIR, will submit a Marie Curie Host Fellowships for Early Stage Training application to call FP6-2002-Mobility-2. If LEAR receives a PhD grant from this (which is by no means certain, even if the application is accepted), the subject will not overlap with PASCAL-RTN themes.

Partner 3, NCRG Aston. The Neural Computing Research Group at Aston will submit a proposal within *Marie Curie Host Fellowship for Early Stage Training* (EST) action. There will be no overlap with the present research project.

Partner 4, LIA Avignon. We are in the final month of the project SMADA (IST-1999-10667), to which we contributed 40 person months.

Partner 5, UPC Barcelona. Dr Jason Weston, currently at the Max-Planck Institute at Tübingen, has submitted an individual application for a Marie-Curie grant to visit José L. Balcázar at UPC for one year, as a postdoc.

Partner 6, IDA Fraunhofer. We are also applying for a grant from another Marie Curie Research Training Network application, STALKER. However there would not be overlap since the topic will be different (the topic is in theory of classification using concentration inequalities).

Partner 7, U Bayreuth. The research group is involved in the MACSI-net IST-1999-14077. There is no overlap with present proposal.

Partner 8, EURANDOM. EURANDOM has one FP5 Marie Curie Individual Fellow program in another area of research (stochastic Operations Research) and has one current application for a Marie Curie Intra European Fellow (FP6), also in another area (probability theory). Furthermore EURANDOM is taking part in two applications for Integrated Projects, both in an area where no clashes with this proposal are to be expected.

Partner 9, TU Graz. We have no similar or overlapping projects or applications.

Partner 10, CoSCo Helsinki. We have no similar or overlapping projects or applications.

Partner 11, MU Leoben. We have no similar or overlapping projects or applications.

Partner 12, JSI Ljubljana. Related NoE proposals within FP6 are PASCAL and KD-Net where we appear as core partners. A related FP5 project was Sol-Eu-Net IST-1999-11495 (2000-2002) on establishing virtual enterprise for data-mining technology. None of these overlap with the current application.

Partner 13, RRG Oxford. The group is also a partner in the Marie-Curie RTN proposal VISIONET-TRAIN, but there is no overlap in the activities involved.

Partner 14, UJM Saint-Etienne. We have no similar or overlapping projects or applications.

Partner 15, I3S Sophia Antipolis. We have no similar or overlapping projects or applications.

Partner 16, KTH Stockholm.

Partner 17, MPI Tübingen. We are also applying for a grant from another Marie Curie Research Training Network application, STALKER, but there will be no overlap as the topic will be different (the theory of classification using concentration inequalities). We are also part of a NoE called Medic.Ace, which aims to develop advanced methods of information processing for medical and health data processing and prediction. Again there is no overlap as this is on a different topic.

Partner 18, U Aalborg. We have no similar or overlapping projects or applications.

Partner 19, UC London. We have no similar or overlapping projects or applications.

Partner 20, U Milan. We have no similar or overlapping projects or applications.

Partner 21, UPMC Paris. We have no similar or overlapping projects or applications.

Partner 22, ETH Zurich. We have no similar or overlapping projects or applications.

B9. Other Issues

B9.1. Information Required from Proposers on the Ethical Aspects of the Presented Research

As far as we know, this work does not raise any particular ethical or safety related issues. It does not involve biological or genetic issues in any way. The research itself does not involve any use of personal data, but if successful, it will simplify data mining and improve data interpretation in many text and multimedia databases, including ones that may include personal data. However we believe that current access legislation is adequate to cover this and that no special attention needs to be given to this issue during the project.

A. Ethical Issues Table

Does the research presented in this proposal raise sensitive ethical questions related to:	YES/NO
Human beings	No
Human biological samples	No
Personal data (whether identified by name or not)	No
Genetic information	No
Animals	No

B. Requested Confirmations

We confirm that this proposal does not in any way involve:

- Research activity aimed at human cloning for reproductive purposes;
- Research activity intended to modify the genetic heritage of human beings which could make such changes heritable;
- Research activity intended to create human embryos solely for the purpose of research or for the purpose of stem cell procurement, including by means of somatic cell nuclear transfer;
- Research involving the use of human embryos or embryonic stem cells with the exception of banked or isolated human embryonic stem cells in culture.

ENDPAGE

HUMAN RESOURCES AND MOBILITY (HRM) ACTIVITIES

MARIE CURIE ACTIONS
Research Training Networks (RTNs)

PART B

PASCAL-RTN

Marie Curie Research Training Network in

Pattern Analysis, Statistical Modelling and Computational Learning

April 2, 2003