Proposal description: STIPCO

1a. Research topic

One of the main trends of statistical physics in the last two decades has been the emergence of new concepts and techniques to study the cooperative behaviour of disordered and frustrated systems and provide the means for their fruitful application. These are systems where the elementary microscopic objects (whether they are for instance spins or particles or human beings), while usually similar to one another in isolation exhibit complex cooperative behaviour and acquire distinction in their individual compromises with their fellows and their environments. This evolution, augmented and intertwined with the progress in the study of out-of-equilibrium systems, has defined new areas where statistical physics has great potential for very fruitful application and has demonstrated this utility. Additionally, there has been significant complementary stimulation of statistical physics through these mathematical and conceptual analogies and extensions, leading to a symbiotic amplification of progress throughout. This network is born precisely to promote the use of statistical physics to study complex systems in realms that are at the borderline of physics.

We have identified a number of research areas outside the traditional frontiers of physics where the application of these methods looks particularly promising, where we have relevant preliminary and complementary expertise and a collaborative network will have a large impact. They all relate in some way to the collective behavior of heterogeneous agents, where “agents” can be as diverse as logical constraints in optimization problems or error correcting codes, proteins in molecular networks, or economic agents. The coherence of our approach derives from the concepts and techniques which are used and from the types of questions that we will address in all these different topics. Our fields of research fall into four main groups:

Combinatorial optimization problems. Hard combinatorial optimization are naturally related to ubiquitous concepts and to questions that are typical of statistical physics, like for example rugged energy landscapes (where one seeks minima), dynamics in these landscapes, thermalization (to provide probabalistic overcoming of barriers) and out-of-equilibrium processes, ... The notions of phase transition and critical phenomena have been found to be of basic relevance in understanding the onset of computational complexity in hard combinatorial optimization problems. The out-of-equilibrium behavior of randomized optimization algorithms is connected to many interesting questions related to statistical physics, such as the existence and the characterization of asymptotic stationary states, and the question of relaxation times.

Error correcting codes. This is a field of significant practical importance, as these codes provide mechanisms for retrieving an original message after corruption due to noise during transmission. They are being used extensively in most means of information transmission, ranging from satellite communication to the storage of information on hardware devices. The observation that error correcting codes are mathematically equivalent to certain theoretical models of spin glasses makes it possible to use modern methods of the theory of disordered systems to obtain new results for error correcting codes. This field is also connected to the combinatorial optimization problems, both from the point of view of some fundamental questions (like estimating the capacity), and of the practical aspect of decoding, and coding optimization.

Biological systems. Biological systems are a source of interesting and relevant problems that can be successfully treated by methods of statistical physics. The recent accumulation of data on genetic sequence and on gene expression signals the existence of complex *correlations* in these systems, and calls for the development of new methods of analysis and classification. In particular, the experience obtained in the study of the statistical physics of disordered systems can be used to gain insight on the working of complex biological *networks*, on the molecular (proteic), genetical and ecological levels. This will need the development of new conceptual tools, since biological systems are strongly out of equilibrium and exhibit a complex spatio-temporal pattern of many interacting units. We will develop such concepts in order to gain a more quantitative understanding of correlations and networks in biology.

Interacting agents in financial markets and socio-economic systems. In the last few years a new research activity has been gathering importance among scientists in theoretical physics. Physicists, especially those working in statistical mechanics, complexity and biophysics, are realizing that ideas and tools of their discipline are of relevance for a range of topics much larger than that identified by the traditional boundaries of physics: the examples that are most relevant here range from economics and financial markets to vehicular traffic and the structure of Internet. Let us stress the difference: while most of the modeling in economy deals with representative agents, in the approach inspired from the statistical physics one considers a large set of *heterogeneous* agents, where everyone acts according to his own set of allowed strategies, normally with finite foresight, but with collective consequence and resultant correlation. When taking this point of view one deals with a much richer situation, where fundamental concepts such as irrational behavior or macroscopic equilibrium acquire a new meaning.

Together with these "technical" reasons (methods of physics can be useful in studying problems related to markets and economics) there is also a paradigmatic shift: more and more physicists today feel their professional responsibility in society undergoing unprecedented changes. The traditional approach based on the study of a single well defined domain no longer meets the complex nature of today's research and society's needs. In the interdisciplinary approach to most of the topics which we have chosen to study, Europeans stands ahead of their US colleagues. We present this network also as a tool to sustain this lead and to maximize unique synergies: we intend to take advantage of this new trend by organizing an Europe-wide collaboration project to tackle challenging tasks in a large domain of natural sciences.

1b. Project objectives

This project is original in the sense that it deals with several different problems, in various fields, where modern statistical physics offers a unifying point of view.

The principal motivation for this network is to provide a means to train young researchers with an appropriate expertise and strong background in statistical physics, disordered systems and out-of-equilibrium dynamics such that they will be able to tackle precisely, accurately, productively and usefully the further diverse problems of collective behaviour of complex cooperative systems which characterize our world in many fields of endeavour. The specific but diverse problems tackled and the cross-disciplinary procedures used in their individual projects will provide a deep training in many intertwined general techniques and concepts, as well as relevant parts of discrete mathematics, economics and biology. Already the cross-disciplinary potential of such studies is recognized, for example in the employment of physicists in finance, and in such fields as neural networks, operational research, error-correction and bioinformatics, but the potential is much greater than the present application, and we believe that the network will provide a unique training opportunity.

To simplify and shorten the presentation we shall list separately the different scientific objectives that we have in mind, organized in four different themes as above, and we will try to stress the common relevant features.

Combinatorial optimization problems.

An example of the relationship between open problems that are of interest both in computer science and in statistical physics is provided by *threshold phenomena*. These are well known from the pioneering works of Erdos and Renyi on random graphs in the '50s, and they have recently attracted a lot of attention in the context of random combinatorial problems. Graph coloring, the satisfiability of Boolean formulae, the vertex covering problem are examples of hard combinatorial problems that may exhibit phase transitions when control parameters, defining distributions of random instances, are modified. Following this line of research we intend to develop a systematic study of the relationship between phase transition, critical phenomena and the onset of computational complexity in hard combinatorial optimization problems. We shall also study the geometrical structure of the space of solutions (quasi-optimal configurations) in optimization problems, and understand the relationship between this structure and the resolution time and memory requirements. We will also be involved with a dynamical analysis: we shall analyze the behavior and the performances of search optimization algorithms, with a particular effort towards stochastic algorithms, for which statistical physics may offer some particularly efficient approaches.

The hardest combinatorial optimization problems of computer science appear to share characteristics that are largely independent from the specific algorithms adopted for their solution. Important features such as the solution time are conjectured to be tightly related to the geometrical structure of their phase space.

Error correcting codes.

This field has recently experienced a true revolution. The first crucial moment has been the invention of a new class of codes (called "turbo codes") by Berrou and Glavieux in 1993, the second has been the rediscovery of "Low Density Parity Check Codes" or LDPC (first discovered by Gallager in 1962, then forgotten and rediscovered in 1997). Both these families contain codes which are much better than anything known before: both of them provide codes which are not very difficult to decode and that approach very closely the Shannon theoretical limit. No such codes were known before. Despite the lack of a good theoretical understanding, turbo codes have already been adopted as the new international standard of communications. The invention of these codes was empirical: there was no theoretical understanding of why they are so successful. The decoding algorithms are heuristic algorithms and they are thought to be approximate.

We have already mapped both turbo codes and LDPC into disordered spin models. This relates the existence of a threshold Θ (such that for a signal to noise ratio larger than Θ it is possible to communicate in an error free manner) to the existence of a phase transition in the related spin model.

One important open issue which we will address concerns the finite size effects, which may be large in the vicinity of the threshold, and that are extremely important for practical applications. We shall also analyze the use of approximations based on statistical physics methods to the decoding problem, and we will study the dynamical behavior of the decoding algorithm (it has been observed numerically that decoding takes much longer time as one approaches to the phase transition, presenting some similarity to the critical slowing down). It can be shown that the decoding algorithm searches for a minimum of an effective free energy: it is not known whether this minimum is unique, but some preliminary indications point to the existence of multiple minima, in some region of the space of parameters. We shall investigate the possible existence of a glassy phase in these systems, which again would be very important for applications, and has not yet being considered in coding. Finally we shall also try to optimize irregular LDPC constructions.

Biological systems.

The principal objective of this collaboration is the study of correlations and networks in biological systems. We will be mainly interested in three different aspects: the analysis of biological data, the dynamical modeling of biological networks and, most importantly, the comparison between data and models.

As far as data analysis is concerned we intend to use cluster analysis to address three different problems. We will first study the data obtained in DNA chips for a large number of genes in a given sample, in order to identify genes with correlated expression levels, to search for genetic tools for advanced diagnosis, prognosis and assessment of the effects of therapy, and to look for clues for the functions of new genes. Another application of cluster analysis will be to identify co–regulated genes, as well as significant regulatory sequences using the large data bases available. We will also improve the identification of proteins with the same fold and similar function, in the difficult case in which they have very low sequence identity.

The modeling of biological networks has three main objectives. We will study dynamical models of evolution, from the point of view of quasi-species evolution in mean-field fitness landscapes, in order to get a better understanding both of the steady state and of the dynamics. Concerning the coevolution of many species, we will develop a 'minimal' phenotypical model for the coevolution of a number of species that can be compared with phylogenetic data and contains the fundamental biological dependencies (e.g., prey-predator relationships). The selective pressures of co-evolution generates networks that are far from randomly connected. We aim at understanding the statistics of their structure as a consequence of the underlying dynamics.

An important aspect of our project is to link the various studies involving dynamics, network structure, and genomic data. Specifically, we have three main objectives in this direction. First is the study of the shape of phylogenetic trees: we will define proper observables that are characteristic of the evolutionary dynamics and at the same time accessible from existing data. Second is the study of the architecture of protein and genetical networks in the cell: we will try to identify the type of architecture from existing data on the number of regulatory proteins in a network. Third is the study of correlations between different parts of the genome, using the fact that the error rate per base pair and per generation of the replication process is inversely proportional to the genome size.

Interacting agents in financial markets and socio-economic systems.

Here again, correlations and competing interactions play a crucial role, but there are also several fundamental differences from conventional physics problems, in that influence is often via commonly available indices and information, rather than direct, information used or anticipated is often highly non-local in time, including anticipations of the future, and non-equilibration is the norm of real systems. In order for the statistical physics to contribute in clarifying important issues concerning financial markets and socio-economic systems, we believe that the most important initial steps are the identification of (i) relevant 'stylized facts' in empirical data (for example, the well established fat tails of price change distributions and long-ranged volatility correlations in financial time series) and (ii) simple 'benchmark' models. These should contain the main relevant ingredients of the problem, in a simplified situation which allows to understand some of the most important mechanisms at work.

The *Minority Game* is an example of such a model, which has recently been invented by some members of our network, and is identified by the community as one such paradigm model for the study of financial markets. This model (and its current variants), despite a simple and highly idealized structure already exhibits complex cooperative behaviour, showing phase transitions, regions of equilibration and non-equilibration, and relevant stochasticity. A combination of simulational, analytic and numerical solution for both static and dynamical properties has exposed several novel features and provided both understanding and procedures to explore deeper and more realistic scenarios.

One main objective is to develop this type of interacting agent models in order to incorporate gradually a more realistic description of the complexity of real markets. Among the directions for extension are the inclusion of variation among agents, of inclination, reliability and frequency of play, as well as of strategies, the inclusion of consumers as well as speculators, of the effects of capital and of speads, of driving forces and predictions and insurances of future actions, finite-size and non-linear effects, rare fluctuations and external shocks, the role of money. One particularly interesting question is the importance of 'market microstructure' effects (price fixing mechanisms) for the collective, large scale properties of a market and its global stability.

Building progressively on these simple models will serve as a guide to define interesting observables (that have not been considered before in the economic or econometric literature), it will allow one to create faithful artificial markets or economies, with the perspective, further down the line, of being able to *simulate* real situations (such as, for example, the influence of the Tobin tax on transactions on the behaviour of currency markets).

We also intend to investigate some other important socio-economic problems like Vehicular traffic in urban road networks (on which important work has already been done), information transfers on the Internet and the dynamics of scientific research development and interaction, using the same blends of techniques, minimalist modelling, simulations, and all the tools of statistical mechanics of disordered systems.

2. Scientific originality of the project

Our project has two main characteristics. It lies at the forefront of research in statistical physics, and it is by nature interdisciplinary. Each of the four directions of research which we have selected represents a very active research field, which has recently seen important developments. The members of the network have played key roles in those development, and in several cases they have even started the whole field (e.g. the application of statistical physics to combinatorial optimization, error correcting codes, or financial and economics problems). Here are the contributions which we expect from this collaboration in each field.

Combinatorial optimization problems.

We aim first at developing, and then at transfering knowledge and techniques from physics to computer science and optimization problems. As an example we remind that phase transitions have been intensively studied in statistical physics for more than a century, and a large bunch of conceptual and technical tools have been developed to deal with them. That such tools can be successfully applied to random combinatorial problems to tackle, and reach some intuition on unclear aspects of threshold phenomena is now established. There is no doubt that random combinatorial problems are well defined mathematical objects and, as such, should ultimately be discussed and understood rigorously. It is however our belief that physical approaches propose new and stimulating perspectives on these issues, that could eventually benefit to mathematicians and computer scientists.

Besides its intrinsic interest, the statistical physics of random combinatorial problems is a lively and true example of an interdisciplinary field of research. The richness and variety of the current research undoubtedly shows that interdisciplinary approaches are not only fascinating but also fructuous.

Error correcting codes.

The relation between error-correcting codes and Ising spin models have first been studied by Sourlas, a member of the ENS node (1989), by mapping LDPC codes onto the Hamiltonian of a multi-spin-interaction model. This work has been recently extended by members of the network to study more general practical Gallager error-correcting codes as well as convolution codes such as the celebrated turbo code.

Statistical physics offers a very efficient tool for the theoretical understanding of various aspects of those efficient codes, which often map to either fully connected, or to one dimensional spin glass like models for which we have very good analytic methods available. Using these methods one can calculate typical properties of specific codes, in contrast to the rigorous general bounds derived in the information theory literature. The latter tend to be relatively loose, and show similarity to certain known approximations used in statistical physics. We are thus confident that this project will make major contributions, particularly in optimization of irregular Gallager codes, in the study of finite size effects (allowing to assess the usefulness of various codes), in the performance evaluation of various decoding dynamics. These are all problems which are both highly challenging theoretical problems, but could also have important practical implications.

Biological systems.

Our project addresses several questions of biology using concepts and tools from statistical physics, very connected to the ones used for solving many of the problems discussed in the other sections. These questions include some long lasting problems which we will study using new methods, as well as some problems which have emerged recently.

We have developed a new method to mine data that come in the form of an array such as e.g. $A_{g,s}$, giving the expression level of gene g in sample s. This methodology was demonstrated to work well for gene expression data, and has been able even to throw new light on old and difficult problems like the spin glass one. We believe that by exporting this novel methodology to additional new problem areas, considerable insight will be gained.

While there exists a large number of studies of molecular and trophic networks and of evolutionary trees, they aim at the reconstruction of a *particular* network or tree. Our project has a different perspective: we want to identify *statistical* features of networks and trees and relate them to stochastic models of evolution. Clearly, this requires a different kind of theoretical models that have 'minimal' interactions and sufficiently coarse-grained degrees of freedom so as to allow for a statistical analysis.

The present treatments to quasi-species evolution achieve the required dimensionality reduction by projecting it to error-class dynamics—what can only be done in very special, highly idealized cases. We build on a statistical mechanics approach to project the equation to a dynamical system in the finite-dimensional space of observable, a concept that is new in this context. The correlation between number of regulatory proteins, mutation rate and network architecture has not yet been studied according to the best of our knowledge.

Based on a comparison of specific models and genomic data, we will be able to gain quantitative insight in the important biological concepts of *evolvability* and *robustness* of co-regulated processes, which, despite their ubiquity, have until now been addressed mainly in a qualitative way. Clearly, this requires a sustained coordination of data-based and theoretical work, which we see as the "raison-d'être" of this collaboration.

Interacting agents in financial markets and socio-economic systems.

The idea of an interdisciplinary endeavors of physicists towards socio-economic sciences is not new. Nobel prize winner physicists Phil Anderson and Murray Gell-Mann have pioneered interdisciplinary endeavors of physicists towards social sciences and economic theory in particular, with the establishment of the Santa Fe Institute. They have organised several interdisciplinary programmes on statistical physics and theoretical economy, which are generally considered as very successful.

Our aim is to promote the advancement of our basic understanding of cornerstone models of complex adaptive systems by exploiting a strong competence in a common theoretical background. The recent development of studies in the minority game, achieved by members of different groups in this network, provides an example of how successful such a strategy can be. This will not only advance our understanding of socio-economic systems but also help developing even further statistical physics, through the new types of questions which are asked. Within this approach, the interaction with experts of other disciplines occurs at the level of the definition of models and at the level of the results. The exchange with experts in other disciplines (economists and social scientists), the confrontation of the models, their results, and available data, will be provided by the development, through this network, of some collaborations which already exist (in the node of Trieste, Orsay, Fribourg ...) and in the interdisciplinary events which we will organize.

3. Research method

Because of its interdisciplinary nature, this project involves a variety of methods, which all relate to the experience and expertise of the research teams.

The main framework to be used is the theory of statistical mechanics, particularly within the context of disordered systems. Statistical mechanics provides the tools to calculate the properties of infinitely large networks exactly, both from the point of view of equilibrium and from that of out of equilibrium dynamics. In addition, results for typical or most probable situations can be derived by averaging over all data realizations for a given structure. In optimisation problems where one typically searches the minimum of a Hamiltonian with quenched disorder, as well as in error correcting codes, which naturally map to random spin systems, these techniques appear as choice methods. They will also be useful in our modelling of socio-economic systems, since they allow to study the dynamical properties of interacting and heterogeneous agents, as well as to part of the biological theme: the networks studied have remarkable common features that call for the methodology of statistical physics: (i) They present a complex spatio-temporal pattern of many interacting units. (ii) They result from a strongly non-equilibrium dynamical process, involving very different time scales, and thus leading to effective quenched disorder. (iii) Most importantly, the recent accumulation of genomic data makes it possible to ask statistical questions that had been out of experimental reach until now.

We have identified a specific set of tools that we will use in these investigations. For each such method, we will indicate below the fields of our project, as described above, to which it applies, with the following code - (a):Combinatorial optimization problems, (b):Error correcting codes, (c):Biological systems, (d): Interacting agents in financial markets and socio-economic systems-. We first list the most important methods which will be used in several of our research themes, providing the links between them. We want to point out that for each of the main methods which we will use, we have in this network some of the leading world experts.

General statistical physics methods. Mean field, variational, renormalization group method are useful in all the fields which we study.

The replica method. This analytic method for averaging over the disorder, which has been developed in the last two decades, will play a central role in most calculations in (a,b), as well as in some of the studies of (d).

Dynamical field theories. The type of methods developed after the original ideas of Martin Siggia Rose is the natural complement of the previous one, which allows to study the dynamical properties in disordered systems. It will be useful in all four themes.

Cavity techniques. In some scenarios the replica technique may prove difficult and will have to be replaced by other methods and/or approximations. One of the methods that may be very handy is the cavity method. This is particular true for the finite connectivity disordered systems appearing naturally in (a,b).

Superparamagnetic Clustering of Data (SPC). This is a powerful technique which we have introduced recently. It uses the physical properties of granular ferromagnets to “solve” the clustering problem. It shows a number of advantages with respect to traditional algorithms (the number of clusters is determined by the algorithm itself, it presents good stability against noise, and is able to identify robust, stable clusters, of possibly irregular, non-spherical shape). It obviously plays a central role in the data analysis part of (c), but it will also be useful for (a) and (b) in order to study the structure of the space of quasi-optimal configurations.

TAP and naive mean-field. In many cases, where the solutions to learning, optimization and decoding problems are difficult to obtain, or require a prohibitively long time, one resorts to approximation techniques which may provide a close-to-optimal solution in a very short time. Naive mean field factorizable representations as well as the more elaborate mean field approaches à la Thouless-Anderson Palmer will be used in (a,b).

Methods of information theory Most of the problems targeted are interdisciplinary and may therefore require the use of completely different theoretical tools used in other disciplines. More specifically, we may resort to methods of information theory, in particular in order to provide bounds on various quantities, in all four themes.

Numerical studies Last but not least! Our analytic studies will be supported by extensive numerical simulations. This is important for two reasons: 1. Analytic results can be validated through simulations, especially when various assumptions are used, as is often the case in analytic studies of disordered systems. 2. This allows for the study of more complex and more realistic models. Since we are typically interested in the behavior of large systems, we need the tools of finite size scaling to obtain properties of infinite networks from small system simulations.

On each theme, we shall also use some specific and specialized techniques, which are listed below.

Coupled Two-Way Clustering (CTWC) is a powerful novel method, used to mine gene expression data in (c). One clusters the samples and then the genes in an iterative and coupled fashion, which breaks down the total data set into subsets of genes and samples that can reveal significant partitions into clusters.

Bayesian methods. The Bayesian approach provides an efficient and useful method for making decisions in random environments by calculating expectation values and error-bar of observables. It is somewhat complementary to the statistical-mechanics based approaches and have been widely employed over the years for a variety of tasks. We will make connections between Bayesian approaches and the results obtained using the methods of statistical physics in (b,c).

Graphical models Decoding in error-correcting codes is typically carried out using methods derived from simple graphical models, such as belief-propagation. We will make use of these methods in our studies and expose existing links between them and the methods of statistical physics (e.g. iteration of Bethe equations). This applies mainly to (b) and (c), but we will also explore the possibility of useful applications to (a).

Statistical field theory of non-equilibrium systems of the reaction-diffusion type. This powerful method, developed in recent years, will be used in (c) in order to study the co-evolution of large number of species, described as some stochastic generalizations of the Lotka-Volterra equations.

4. Work plan

Combinatorial optimisation problems. The first two years will be dedicated to the study of the relationship between phase transitions, critical phenomena and the onset of computational complexity in hard combinatorial optimization problems (collaboration Trieste, ENS, Orsay, Barcelona, Oxford), as well as the geometrical structure of the space of solutions (Trieste, ENS, Barcelona, Weizmann).

In the third and fourth year, we shall turn to the study of dynamical problems, particularly the numerical and analytic study of the performances of search optimization algorithms. We shall also study in this second phase the combinatorial decision problems under uncertainty and stochastic planning.

This part research will be headed by the groups in Trieste and ENS, which will be in charge of coordinating the contributions of the other groups and developing the collaboration.

Error correcting codes. The first two years will be mostly dedicated to developing the methods and models to be used in this project, as well as to studying the approaches used within the information theory community. In particular, we will focus on the development of variational methods aimed at optimising Low Density Parity Check (LDPC) code constructions and on the weight enumerator and reliability function techniques for evaluating the performance degradation below and above Shannon's bound. This will be complemented by other numerical and analytical studies of finite size effects. The phase diagram should also be established by the end of the second year.

In the third year we will concentrate on studying various decoding methods for improving the performance of both LDPC and generalised convolutional codes. The decoding methods studied will be derived mainly from existing approximation methods used in the statistical physics community, and will be examined with emphasis on improving the speed and efficiency of the decoding process in specific codes and constructions.

The fourth year will be mainly dedicated to establishing links between information theory and statistical physics in a variety of problems. We will take advantage of the typical exact results which can be calculated for specific codes using the methods of statistical physics, to go beyond the general bounds provided in the information theory literature.

This research will be done mainly in ENS and Aston, who will coordinate it, with contributions from Rome, Orsay and Trieste.

Statistical physics of biological systems. In the first two years of the program, the data analysis and the theoretical work will be done in parallel. Priority will be given to those areas where the current state of genomic data seems most promising. We expect substantial results by the end of the second year on the identification of genes with correlated expression levels and the co–regulated genes. We will also have developed at that time the 'minimal' phenotypical model for the coevolution of several species. The first two phases of the work on trophic and molecular networks will have been completed: (i) The statistical analysis of existing networks and phylogenies, (ii) The development of observables that characterize these networks in a quantitative way.

In the third and fourth year, we plan an intensive feedback between theory and data analysis. They will be dedicated to the influence of genome size and the number of regulatory proteins on network architecture, as well as the development of dynamical models that lead to networks and phylogenies, and the quantitative comparison between the model and the data.

The coordinating nodes for this topic are Koeln, Roma and Weizmann, with contributions coming from Barcelona and Orsay.

Interacting agents in financial markets and socio-economic systems. In the first two years we shall generalize the models of heterogeneous interacting agents (minority Games) for financial markets in many different directions. We will consider agents who have diverse investment strategies on a set of assets, or, very importantly, over different time scales. Intertemporal aspects of financial time series are among the best characterized 'stylized facts' that need to be understood. By the end of the second year we expect to have realistic such generalisations, which take into account relevant economic or financial constraints (wealth balance, market clearing, etc.)

In the third and fourth year we shall characterize the various phases of these interacting agent systems, and the corresponding phenomenology of the artificial markets/economy that these models aim at describing. Based on our understanding of simple models we will pursue the developement of empirical measures of informational efficiency and of risk for financial markets. These will be compared with empirical data on real financial markets. We will study the impact of market microstructure (price forming mechanism) on the large scale properties of these markets and their stability.

We shall also develop other models for socio-economic systems such as information networks on the internet, urban traffic management, social choice.

The nodes which will coordinate this activity, are Trieste, Orsay and Oxford. This activity will also involve the nodes in Fribourg and Rome.

Professional research effort on the network project			
Participant	Young researchers to be financed by the contract (person-months) (a)	Researchers to be financed from other sources (person-months) (b)	Researchers likely to contribute to the project (number of individuals) (c)
1.	36	108	6
2.	24	50	6
3.	24	100	4
4.	20	43	4
5.	18	43	5
6.	24	134	6
7.	34	148	9
8.	36	168	7
9.	36	175	10
10.	23	46	4
Totals	275	1015	60

5. Collective expertise

NODE 1: ORSAY

The Orsay node consists of two groups, one is in the Laboratoire de Physique Theorique et Modeles Statistiques in Orsay, the other is in the Service de Physique de l'Etat Condense in Saclay. The heads of the two groups, M. Mézard and J.-P. Bouchaud, have often collaborated in the past, and will work jointly on this project. The neighborhood of the two labs, which are a couple of kilometers apart, allow for the post-docs to interact simultaneously with researchers in the two groups. This node is recognised internationally as a leading team in the general study of disordered systems, and in particular in the application of statistical physics to the fields of optimization and econophysics. It has also contributed fruitfully to the statistical physics of biomolecules.

This group will be active mainly in three of the themes. On the side of optimization problems, Marc Mézard has a long record, having written with Vannimenus what is probably the first paper on the use of statistical physics to study analytically random optimization problems (the travelling salesman problem-TSP), and having then worked out the solution of the random link matching and TSP. His recent work on the replica symmetry breaking solution of dilute spin glass models opens the way to systematic studies of various combinatorial optimization problems, starting with the K-satisfiability problem; this same method should also be applicable to the study of the glass phase in error correcting codes. Another aspect of our expertise which will be useful in these two themes is the use of mean field equations to generate new algorithms for optimizing and decoding. Interactions on all these themes already exist with the nodes in Paris and Trieste, and certainly we expect to develop tighter links with other goups in this network (Aston, Barcelona, Oxford).

On the side of econophysics, beside the well recognised experience of J.-P. Bouchaud on finance, this node has moved recently towards the construction of various models of interacing agents to study some situations of economy. This has led for instance to the discovery of a generic mechanism for the generation of Pareto (power law) distributions of wealth and to some modelling of volatility clustering. This network will provide the opportunity to develop some tight interactions with the nodes in London, Trieste, and Fribourg on these topics.

Key scientific staff: Dr M. Mézard (50%), Dr J.-P. Bouchaud (50%), Dr I. Giardina (50%), M. Müeller (75%).

Two significant recent publications:

"Wealth condensation in a simple model of economy", J.-P. Bouchaud and M. Mézard, Physica A, **282**, 536 (2000).

“On a universal mechanism for long ranged volatility correlations” Jean-Philippe Bouchaud, Irene Giardina , Marc Mézard, to appear in Quantitative Finance **1**, 212 (2001).

NODE 2:ASTON

The Aston Neural Computing Research Group is widely regarded as one of the most significant and focussed university group in neural computing and related areas in Europe, and consists of two Professors, one Reader, four Lecturers, six postdoctoral researchers, support staff, and students. The group is strong in both theoretical and practical areas and in addition to its high international research profile, has strong interactions with several industries directly and with organisations (such as the Natural Computing Applications Forum) actively attempting to strengthen the links between academia and industry.

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.

Research activities in recent years concentrated on theoretical analysis of Low Density Parity Check codes, matrix based public-key cryptosystems and advanced methods of pattern analysis, mainly support vector machines and Gaussian Processes; the analysis has been carried out using tools adopted from statistical physics. As part of this activity we analysed both Gallager and MN codes, identifying their dependence on the properties of the construction used. In addition, we studied certain practical decoding methods and their similarity to existing methods of statistical physics. Our expertise and competence in the fields of error-correcting codes, pattern analysis and in the link between statistical physics and information theory make our group particularly suited for carrying out further research in this area as planned in this network.

The NCRG will contribute to this project through direct collaboration with other teams in the network. For example, the groups of both Aston and Paris have been working on similar aspects of error-correcting codes in the last 3-4 years, it is only natural that a collaboration between the two groups will flourish within the network. In addition, one can point to several links between the models we examine and other problems studied in this project, such as combinatorial optimisation and bioinformatics.

Key scientific staff: Prof. D. Saad (20%), Dr. M. Opper (20%), Dr. D. Malzhan (15%), Dr. J. van-Mourik (25%) and Dr. N. Skantos (25%).

Two significant recent publications

Y. Kabashima, T. Murayama and D. Saad, "Typical Performance of Gallager-type Error-Correcting Codes", *Phys. Rev. Lett.* **84**, 1355-1358, (2000).

R. Dietrich, M. Opper, H. Sompolinsky "Statistical Mechanics of Support Vector Networks" *Phys. Rev. Lett.*, **82**, 2975-2978, (1999).

NODE 3: BARCELONA

The Barcelona team is very well recognised for its good expertise in computer simulation techniques (finite-temperature Monte Carlo methods, improved algorithms) for the analysis of rugged free energy landscapes in frustrated and disordered systems in general. Conrad Perez-Vicente has introduced neural network models with low activity and Felix Ritort has proposed new models for the study of slow dynamics in glassy systems. This perfectly links with the good theoretical knowledge and research lines undertaken by the other teams of the network. Our activity will be focused onto the study of optimization problems (phase diagrams, algorithm performance) and the statistical physics of biological systems (data analysis, network studies).

Concerning previous contacts, the Barcelona team has seen in the past a good and fruitful collaboration with the teams of Oxford, Rome and Trieste which have lead to several scientific publications in the field of disordered systems and glassy dynamics. Also it has kept contacts and punctual collaborations with the nodes Orsay and Paris in several problems related to the content of the project. This project should help to establish exchanges with the rest of nodes of the network which will lead to new scientific collaborations. Moreover we expect collaborations with some of the nodes (eg. nodes Oxford and Paris) to be particularly strengthened as a result of the collaboration.

Key scientific staff: Prof. F. Ritort (50%), Prof. C. Perez-Vicente (25%), A. A. Garriga (25 %).

Two significant recent publications:

1) L. L. Bonilla, C. Perez-Vicente, F. Ritort and J. Soler, Exactly solvable phase oscillator models with synchornization dynamics, Physical Review Letters **81** (1998) 3643-3646

2) F. Ritort, Solvable dynamics in a system of interacting random tops, Physical Review Letters **80** (1998) 6-9

NODE 4: FRIBOURG

In the recent years more and more physicists have realized that their tools developed for nonlinear and non-equilibrium physics could actually be used to tackle a much wider class of problems, notably in economics. A new branch of science, called econophysics, is emerging, and the Fribourg team has played an important role in this progress. It also maintains an important data base on econophysics which is accessed worldwide.

Our main current contribution has been the introduction and study of the so-called minority game, a stylized model aimed at understanding the market mechanisms at work when a great number of selfish players interact together, like that in a stock market. Interesting collective behavior emerges in this game, which we have been able to study for the first time systematically. Also we have done rather systematic work on risk evaluation using portfolio theory. The "old" Fribourg team now is spread out to Oxford (D. Challet) and to Trieste (M. Marsili), forming the basis for very natural interactions, and some new members (Capocci and Laureti) have joined Fribourg.

Our current work is diversifying towards a few new area: Information Theory used to understand Internet and massive large information network. The importance and timeliness of such studies cannot be overemphasized. So far this area has been tried only in a rather artisanal way. However the web reached an age where much more refined understanding, beyond common sense analysis, is needed. We have been able to identify a few fundamental processes and a prototype model to tackle this problem, and we hope to be able to develop it thanks to this network.

In the past, we have often collaborated with the Oxford, Trieste, and Rome teams. Indeed, some of our main papers were published jointly with members from those teams. Besides, we have much affinity with the Paris, Orsay and Weizmann teams as well. We expect to be able to develop some collaborations with all these groups, particularly on the topic of socio economic systems.

Key scientific staff:

Prof. yi-cheng Zhang 25%, Andrea Capocci 33%, Paolo Laureti 33%.

Two significant recent publications:

" Emergence of Cooperation in an Evolutionary Game", D. Challet and Y.-C. Zhang, Physica A 246 (1997) 407.

" Modeling Market Mechanism with Minority Games", D. Challet, M Marsili and Y-C Zhang Physica A276 (2000) 284-315.

NODE 5: KOELN

Michael Laessig has a strong expertise in the renormalization and nonperturbative techniques applied to the physics of growth processes and to systems with quenched disorder. He and his group have co-initiated the statistical physics studies of sequence similarity, and the Koeln group has experience in the dynamical modeling of co-evolution. In particular, it has recently identified the most relevant types of interactions between species in such networks.

This group will be mostly active on the topic of biological systems, where this accumulated expertise will be very useful. The renormalization techniques will be relevant, in particular, for analyzing spatio-temporal patterns in population systems, and the experience in sequence similarity will be used in quantifying the statistical properties of sequence-based phylogenetic trees.

This group already has some contacts with the nodes in Orsay, Roma/Napoli and Weizmann. In particular, Eytan Domani, Michael Laessig, and Luca Peliti are all involved now in the workshop on Statistical Physics of Biological Information at the University of California (Santa Barbara), where they are spending several months of project preparation and discussion of common interests.

Key scientific staff: Prof. M. Laessig 33 % Dr Martin Rost (Postdoc) 50%, Achmet Atik (PhD Student) 25%, N.N. (PhD Student) 25%.

Two significant recent publications:

D. Drasdo, T. Hwa, and M. Lässig, Scaling laws and similarity detection in sequence alignment with gaps, J. Comput. Biol. 7, 115 (2000).

M. Lässig, U. Bastolla, S. Manrubia, and A. Valleriani, Shape of species networks, Phys. Rev. Lett., to appear (2001).

NODE 6: OXFORD

This node which is based in Theoretical Physics at Oxford and includes also the Neural Networks and Disordered Systems group in the Mathematics Department at Kings College, London, has a broad expertise in many aspects of disordered systems, and in their application to many problems outside of physics. D. Sherrington has been involved actively and fruitfully in the subject of statistical physics of complex systems, both basic and in application, for more than 25 years. Inventor of a fundamental model which was the basis of explosion of interest on disordered systems, he has been involved in many developments since, including early applications to information processing (e.g. neural networks) and hard-combinatorial optimization (e.g. graph partitioning) and recent advances to economics (e.g. thermal minority game). Recently he gave the Royal Society's 2001 Bakerian Lecture, the premier annual prize lecture in the physical sciences, on the subject of this proposal, with the title "Magnets, Microchips, Memories and Markets: Statistical Physics of Complex Systems". He has coordinated several European Union networks and collaborated with several persons from partner countries associated with this application. A.C.C. Coolen has worked for fifteen years in the theory of neural networks, sometimes with D. Sherrington (especially on dynamics away from equilibrium and with different fundamental timescales, work complementary to that of French and Italian partners). He has set up a group on neural networks and disordered systems, and has written a recent important paper on the dynamics of the minority game (including away from equilibrium). D. Challet is the co-inventor of the "minority game" with Zhang (Fribourg), a model for which he has been able to compute equilibrium properties using replica methods, in collaboration with Marsili and Zecchina (Trieste). He also maintains the minority game section of an important data base on econophysics which is accessed worldwide. The team is completed by J.P. Garrahan who has expertise in field theory and statistical physics, and was another of the co-introducers and analysts of the thermal minority game, together with I. Giardina (now node 1) and A. Cavagna, and by E. Moro, an expert on stochastic processes and computer simulations, who has also been crucially involved in the recent studies of economics.

The Oxford team has experience under all four of the main identified activity areas and anticipates this continuing. Several connections are already noted briefly above. Many associations with other nodes already exist and will grow. DS introduced replica symmetry breaking for dilute networks, now further examined in Trieste and Paris. Statistical physics of graph partitioning was initiated jointly with Sourlas of node 7. Recent studies of error correcting codes are relevant to nodes 7 and 2.

Thus Oxford has clear world-class expertise and world-leading experience. All the above scientists are fully active. Also, Theoretical Physics in Oxford (host) has much other expertise in potentially relevant areas of statistical physics, condensed matter and elementary particle physics, as well as complementary biological, mathematical and engineering departments, all of which have received the highest possible (5*) grading in National Research Assessments. Excellent graduate and post-doctoral courses, seminars etc. are available and both students and post-docs attracted to Oxford are of extremely high quality. We have also excellent computing facilities. Oxford's Theoretical Physics is extremely cosmopolitan with members from many countries and with strong research links with many countries, including the groups of this proposal but also many more. Consequently it offers an outstanding training opportunity.

Key scientific staff: Scientific staff: Prof. D. Sherrington (30%), Prof. A.C.C. Coolen (50%), Dr. J.P. Garrahan (50%), Dr. E. Moro (50%), Dr. D. Challet (50%), Dr. L. Berthier (50%).

Two significant recent publications :

"Continuous time dynamics of the thermal minority game", J.P. Garrahan, E. Moro and D. Sherrington, Phys. Rev. E62, R9 (2000)

"Generating functional analysis of the dynamics of the batch minority game with random external information", J.A.F. Heimel and A.C.C. Coolen, cond-mat/0012045 (to be published in Phys. Rev. E).

NODE 7: PARIS

The node in ENS-Paris involves an interdisciplinary team of statistical physicists and theoretical computer scientists, sharing common interests on optimization and information theory (error correcting codes).

L. Cugliandolo, R. Monasson, N. Sourlas, J. Vannimenus have worked on various aspects of disordered statistical systems using theoretical and numerical techniques. Their works have been devoted to spin-glasses, random fields systems, neural networks, out-of-equilibrium dynamics in glassy systems, cellular automata, ... There is a strong tradition of interdisciplinarity at ENS. J. Vannimenus was among the very first physicists, with M. Mezard (Orsay node), to apply stat. mech. concepts and techniques to optimization e.g. the Travelling Salesman Problem in 1984. N. Sourlas has first established in 1989 the mathematical equivalence between error correcting codes and statistical mechanics, and is currently working on the theory of the newly discovered error correcting codes (turbocodes and low density parity check codes). Recently, R. Monasson has pursued this line and studied the existence of phase transition phenoma in optimization problems, and the relationship with computational complexity and initiated the statistical mechanics analysis of backtrack algorithms. L. Cugliandolo, an expert in out-of- equilibrium glassy dynamics, is now focussing on how to apply stat. mech. to the analysis of stochastic search algorithms. The computer science/information theory team is first composed of O. Dubois, who, as a main activity, studies phase transitions in combinatorial problems. His works have permitted to confirm rigorously some results obtained by the physicists in our network, and, by drawing his inspiration directly from the latter, to improve dramatically the performances of algorithmic methods for solving computing problems. Philippe Flajolet is a leading developer in "analytic combinatorics", which is an original blend of combinatorial and complex-asymptotic methods. The role of P. Flajolet is essentially as support and consultant to the physicists of the node in the area of exactly solvable models and rigorous asymptotic methods in random combinatorics. Finally, R. Urbanke is working on the theoretical and applied aspects of information and coding theory, mainly on turbocodes and low density parity check codes, of huge practical importance.

This team has strong links and recent collaborations with the other nodes of the network, on various aspects of the statistical physics of disordered systems. Among these links, let us cite: - the collaboration with the Trieste group: R. Monasson has been collaborating with R. Zecchina, M. Weigt on optimization for six years. N. Sourlas, R. Urbanke are organizing with S. Franz, a workshop on "codes on graphs" and statistical physics, to be held next May. - the collaboration with the Oxford and Rome group, with which L. Cugliandolo, R. Monasson, N. Sourlas have interacted and realized joint works. - the collaboration with the Weizmann group: E. Domany was visiting professor with J. Vannimenus at ENS two years ago for an extended period. - an obvious collaboration with the nearby node of Orsay. This network will be a major opportunity to establish, or strengthen, our links with: - D. Saad, M. Opper (Aston node) on error-correcting codes, - F. Ritort (Barcelona node) on optimization and complex landscapes in disordered systems.

Key scientific staff: Leticia Cugliandolo (Laboratory of Theoretical Physics, ENS,30%), Remi Monasson (Laboratory of Theoretical Physics, ENS,70%), Nicolas Sourlas (Laboratory of Theoretical Physics, ENS, 85 %), Jean Vannimenus (Laboratory of Statistical Physics, ENS, 50 %), Olivier Dubois (LIP6, Computer science, University of Paris 6, 50 %) Philippe Flajolet (INRIA, Computer science, Rocquencourt, 10%), Ruediger Urbanke (Laboratory of communication theory, EPFL, Lausanne, 25%).

Two significant recent publications:

R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, L. Troyansky, "Determining computational complexity from characteristic phase transitions", Nature 400, 133 (1999);

A. Montanari, N. Sourlas, "The statistical mechanics of turbo codes", Eur. Phys. J. B 18, 107 (2000).

NODE 8: ROMA

The group in Rome has a longtime experience on the study of disordered systems and of complexity, both from numerical as well as analytical point of view. Here we will only discuss the most recent of our achievements on this field. There is a deep overlap and synergy with the Naples group, where the main focus is at present on the study of biological systems, and which is part of this node.

In recent years a significant effort has been devoted to understand slow relaxation dynamics observed in many, apparently unrelated, systems such as proteins, disordered and granular materials, spin glasses. In such systems the characteristic relaxation time may change of many orders of magnitude if the external parameters, e.g. the temperature T, are slightly varied. As a consequence correlations display a non-exponential behavior, and equilibration processes slow down giving rise to non-equilibrium phenomena known as *aging*.

The common denominator which makes all these systems displaying similar behavior near the dynamical critical temperature is the *complexity of the free energy landscape*. The trajectory of the representative point in the configuration space can be viewed as a path in a multidimensional potential energy surface. The dynamics is therefore strongly influenced by the complexity of the potential energy landscape: local minima, barriers heights, basin of attraction an other topological properties all influence the dynamics.

The groups in Rome and Naples have been among the first ones in applying disordered system statistical mechanics techniques to the analysis of complex and glassy systems which is now one of the most used ways of investigation, and have an internationally well recognised expertise in this field.

Inside this network, this group will play a major role first in the study of biological systems, but also in the study of optimization problems, where the experience in Rome will be particularly useful to study the phase diagrams and the dynamical performances of the algorithms.

Traditionnally, the Rome group has had many collaborations over the past years with the nodes in ENS, Orsay, Barcelona, Oxford and Weizmann. The network will strengthen these links but also allow to develop new collaborations, in particular the ones with Koeln on biological systems.

Key scientific staff: Prof. Andrea Crisanti (70%), Prof. Enzo Marinari (70%), Prof. Luca Peliti (70%), Dr. Lorenzo Bernardi (70%), Dr. Andrea Rocco (70%).

Two significant recent publications:

Alain Billoire and Enzo Marinari, "Evidences Against Temperature Chaos in Mean Field and Realistic Spin Glasses", J. Phys. A 33 (2000) L265;

Yariv Kafri, David Mukamel and Luca Peliti, "Why is the DNA Denaturation Transition First Order?", Phys. Rev. Lett. 85 (2000) 4988.

NODE 9: TRIESTE

The node in Trieste consists of a rather large team of researchers from the Abdus Salam ICTP and from the SISSA. It is active on all the research lines of this project. Its contributions on protein folding, self-organized criticality, hard combinatorial optimization problems and micro-economical modelling and glassy physics have had high impact in the community. Current research focuses on statistical physics approaches on interdisciplinary research topics ranging from biology to computer science, economics and finance. Our contributions on protein folding, self-organized criticality, hard combinatorial optimization problems and micro-economical modelling and glassy physics have had high impact in the community. M. Marsili, researcher of INFM at SISSA, has a broad expertise on theoretical and numerical methods in statistical physics and stochastic processes. He is actually involved on the front of the statistical-physics approach to economical and game theoretical problems, where he made important contributions on the theoretical study of simple models of financial markets with heterogeneous interacting agents, together with Challet (Oxford), Zhang (Fribourg), Berg and Zecchina. S. Franz is assistant research scientist at ICTP, with a great expertise on disordered systems. His research has focused on glassy systems, with applications to neural networks and models of biological evolution. A. Maritan is full professor at SISSA and has expertise in theoretical statistical physics, a discipline where he gave multiple contributions with emphasis on interdisciplinary aspects. Lately he has been interested to the statistics of river networks and protein folding. M. A. Virasoro is the director of the ICTP; after important studies in high energy physics in the seventies, he devoted himself to statistical physics of disordered systems where he gave fundamental contributions both at a methodological and at a interpretational level to the physics of broken ergodicity of spin glasses and to interdisciplinary applications to neurosciences and to economy. R. Zecchina is assistant research scientist at ICTP, he has competences on many theoretical aspects at the interface between statistical physics and computer science. He has carried research on applications of statistical physics to computer science, neural networks and complexity theory, and has recently solved cornerstone models in combinatorial optimization. Two post docs J. Berg (ICTP), who has expertise on evolutionary game theory, granular systems and microeconomic models, and C. Micheletti (SISSA) who has done research on protein folding, complement this team.

This group has had long lasting collaborations with the nodes in ENS, Orsay and Rome, and many contacts with the all the other teams of the network, thanks to the key role played by the ICTP in organising training sessions on the type of topics studied in this project.

Key scientific staff: Dr J. Berg (50%), Dr S. Franz (50%), M. Leone (50%), Prof. A. Maritan (50%), Prof. M. Marsili (50%), Dr C. Micheletti (50%), Prof. M.A. Virasoro (15%), Dr R. Zecchina(50%).

Two significant recent publications

[1] D. Challet, M. Marsili, R. Zecchina, Phys. Rev. Lett. Vol. 84 pages 1824-1827 (2000)

[2] R. Monasson, R. Zecchina, S. Kirkpatrick B.Selman, L.Troyansky Nature Vol. 400, page 133 (1999).

NODE 10: WEIZMANN

The leader of this node is Eytan Domany who has a well recognised expertise in statistical physics. His recent work has particularly developed towards interdisciplinary research. He was involved in the development of a novel clustering algorithm, SPC, based on the physics of granular magnets, which has proven advantages over most other methods. This algorithm has been sucessfully applied to analyze, among other kinds of information, yeast gene expression data and to identify p53 primary target genes. SPC serves as an important module in a recently developed methodology, Coupled Two Way Clustering of large tables of gene expression data obtained from multiple human samples. The methodology of clustering has been successfully applied to gain insight into more traditional physics problems, such as short range spin glasses, and this network will provide the opportunity to try to "export" it to a wider variety of problem areas within the context of the present proposal, and to develop some collaborations on these topics with other groups in the network.

Many scientific links already exist with other nodes. Eytan Domany has spent a Sabbatical year at Oxford and has collaborated since then with David Sherrington. Luca Peliti (Rome/Napoli) is a frequent visitor at the Weizmann Institute. A collaboration on biological systems has been initiated recently with Marinari (Rome), and application of clustering to the results obtained by Michael Lassig (Koeln) for dynamic evolutionary networks is hoped for.

Key scientific staff: Prof. E. Domany 20% Dr H. Agrawal 50%, Dr D. Volk 25%.

Two significant recent publications

'Coupled two-way clustering analysis of gene microarray data' G. Getz, E. Levine and E. Domany, *PNAS* **97**12079 (2000),

'Spin Domains Generate Hierarchical Ground State Structure in $J = \pm 1$ Spin Glasses' G. Hed, A. K. Hartmann, D.Stauffer and E. Domany, Phys. Rev. Lett. **86**, 3148 (2001).

6. Collaboration

As attested by the section on expertise, the teams involved in this network have a long tradition of collaboration, which have developed through short term exchanges or longer visits, and have resulted in many joint publications.

The research topics of this project have been identified as some top priority topics for us, for which the network, and the collaborations which it will generate, will play a key role.

The collaboration will develop through short term, and occasionally longer term, visits between the different nodes. In particular we intend to favour visits to other nodes of the post docs, who are likely to take the greatest advantage of these exchanges in the sense of developing a true multidisciplinary culture. Although the topics which we will study belong to apparently very different fields, in fact they are conceptually and technically linked via concepts and mathematical formulations, and many of the members of this project are well recognised experts in *several* of our four main fields of research. Our aim is to have each of the post docs develop a top class expertise on one of the topics, but also understand the generality of the approach, see how similar methodologies are used on other topics, and in the end develop the mixture of focused and blended expertise which may lead to breakthroughs in this type of science, in whatever physical, biological, economic or social guise it presents itself. We shall encourage, whenever possible, having a post doc spend a significant part of her/his time in some other nodes (always in a country different from their country of citizenship), or in two laboratories connected to the same node (e.g. Roma/Napoli, Orsay/Saclay, Oxford/Kings College), which should be somewhat easier to achieve. We shall also encourage the exchanges of younger (pre-doc) researchers, even though they are not supported by the network.

We intend to organise a yearly collaboration meeting where the progress on the various research topics will be presented and discussed, but also some very small meetings of 5 to 15 participants for some highly specialised studies.

7. Organisation and management

The network consists of ten nodes. The principal coordinator, Marc Mézard, is also the coordinator of the Orsay node. He will be seconded in his work of principal coordinator by Enzo Marinari (Roma). In each of the other nine nodes, there is one scientist who is responsible for the research carried out in his node. These are: D. Saad in Aston, F. Ritort in Barcelona, Y.-C. Zhang in Fribourg, M. Laessig in Koeln, D. Sherrington in Oxford, R. Monasson in ENS-Paris, E. Marinari in Rome, M. Marsili in Trieste, E. Domany in Weizmann.

These nine local coordinators are in constant contact with the principal coordinator through email, and they will be in charge of deciding and organising the different visits, participations to the meetings, etc..., for their nodes.

In parallel to this organisation, we also have, for each of the four research topics identified on the network, a group of scientists who are in charge of coordinating the scientific activity. These are:

Combinatorial optimization problems: R. Monasson (ENS), R. Zecchina (Trieste)

Error correcting codes: N. Sourlas(ENS) , D. Saad (Aston)

Biological systems: E. Domany (Weizmann), M. Laessig (Koeln), L. Peliti(Roma)

Interacting agents in financial markets and socio-economic systems: J.P. Bouchaud (Orsay), M. Marsili (Trieste), D. Sherrington (Oxford)

All the papers written within this colaboration will be sent to a www archive accessible to all members of the network. We shall also set up a web page of the network, which will describe the activities going on, announce the meetings, advertise the positions available, and have pointers to the papers written by the collaboration. The positions will also be advertised by emails and press announcement.

The principal coordinator has a long management experience, having co-organised many schools and meetings. In particular, he has been the principal coordinator of a Human Capital and Mobility Network.

8. Training need

The research field of this network is a new and rapidly evolving one. While its importance starts to be recognised, it is not yet well represented in the standard education programmes of the universities. Contrarily to some domains of physics which have been well established for a long time, like particle physics for instance, this domain of statistical physics of 'complex' systems is not as well developed in each individual country as to be able to provide a full fledged multidisciplinary training. At the European level of this network, on the contrary, we constitute a sizeable group of well recognised scientists, a group which is able to provide unique opportunities for training, which do not exist anywhere. This is a clear opportunity for Europe to provide the means for a systematic collective expansion in a domain where European scientists, individually, have played already a very important role.

The training that will be provided by this network will select the brightest candidates from our best graduate schools, and offer them the highest level of training in statistical physics, as well as an opening towards several very different domains, an education which will make them both experts in some topic and at the same time versatile.

The recent experience with young researchers having received a similar formation is that they present an enormous potential and easily find jobs in the private industry, in areas such as biotechnologies, finance, systems modelling or various domains of computer science (neural networks, error correcting codes), while some of them also pursue academic careers.

9. Justification of the appointment of young researchers

Young researchers to be financed by the contract				
Participant	Young pre-doctoral researchers to be financed by the contract (person-months) (a)	Young postdoctoral researchers to be financed by the contract (person-months) (b)	Total(a+b) (c)	Scientific specialities in which training will be provided (d)
1.	0	36	36	P-07,S-13,S-22
2.	0	24	24	P-07,M-09
3.	0	24	24	P-07
4.	0	20	20	P-07,S-22
5.	0	18	18	P-07,L-08
6.	0	24	24	P-07,S-13,S-22
7.	0	34	34	P-07,M-06
8.	0	36	36	P-07,L-08,M-05
9.	0	36	36	P-07,M-06
10.	0	23	23	P-07,L-08
Totals	0	275	Overall Total 275	

The larger numbers for nodes 1,7,8,9 are justified by the fact that they all involve several research groups. The typical length of the appointments will be between one and two years. We shall also look for candidates who will be able to spend one post-doc in one of our nodes, followed immediately by another post-doc in another node (both nodes being dstinct from the country of citizenship of the post doc, and from the country where she/he will have had her/his graduate training). The job openings will be advertised widely, in particular they will be posted on our Web site. We shall select the best candidates in their fields, paying attention to provide equal opportunity to all, irrespective of their gender or their country of origin. We shall also attempt to ensure that, whenever practical, both genders are represented on panels making decisions on appointments.

In each node, this network is headed by a scientist with a strong international recognition, as well as a good experience at supervising young researchers. From our past experience, we are confident that the topics of this network and the opportunity of training it provides will be able to attract enough very high level applications.

10. Training programme

In this network, each node will provide training for several graduate students and post-docs, financed by various sources. We are asking for the EC support for young post-docs because we believe that the post-doc period is the one where the young researchers can take the best advantage of a mobility in other European countries. Most of the nodes of the network involve very big and famous universities (or equivalent formation centers) which provide a very wide spectrum of post doctoral lectures and seminars in different disciplines, and we intend that all the post docs whom we will hire will be able to benefit from these.

Beside the training obtained locally at the node where she/he will be appointed, the post doc will receive training through visits and secondments to other nodes of the network, and through participation in the workshops which we will organise. The post-docs will of course be invited to all these events, and we shall strongly encourage them to give talks about their work on these occasions.

We also intend to organise some schools for topical training. This aspect of the network will greatly benefit from the synergy of our project with some of the objectives of the ICTP in Trieste, where several scientists belong to our Trieste node. The ICTP organizes extended visiting scientists and summer schools programs, and has preliminary plans for two such events, on topics at the frontier between statistical physics and other disciplines like computer science, economics and bioinformatics, for the year 2002. These events could provide a full fledged multidisciplinary training and become a concrete meeting point for the network, as well as an opening to the rest of the research community.

In addition to the persons employed on the contact we shall train many other pre- and post-doctoral researchers, supported in other ways, who will benefit from the exposure and association which this network will provide.

11. Multidisciplinarity in the training programme

As we have stressed before, this project is multidisciplinary in nature. In each of our topics, we have contacts with researchers from other communities with whom the post-docs will be able to interact. Here is a partial list of these contacts:

Combinatorial optimization problems. The presence of O. Dubois and P. Flajolet in the network (in the ENS-Paris node) shows the tight links with computer science and combinatorics. We also have contacts with J. Chayes and C. Borg at Microsoft.

Error correcting codes. We have links with R. Urbanke, a great expert of coding theory, who is associatd with the ENS-Paris node, and also with J. Yedidia, who works in practical applications of codes for Mitsubishi. N. Sourlas is co-organising this month a specialised meeting in the ICTP, Trieste, together with F. Forney, another recognised world expert.

Biological systems. The data analysis work performed in the Weizmann is done in tight collaboration with several biological groups. We also have good contacts through our Oxford node with S. Kauffman at the Santa Fe Institute, and also with G. West and H. Frauenfelder (Los Alamos National Laboratory)

Interacting agents in financial markets and socio-economic systems. The groups in Trieste, Oxford and Orsay (Saclay) have established interesting links with researchers in economics and finance: A. Rustichini (Boston), D. Farmer (Santa Fe), T. Lux (Kiel), M. Potters (Paris), H. E. Stanley (Boston) R. Cont (Polytechnique), J. Voit (Bayreuth), E. Aurell (Stockholm).

12. Connections with industry in the training programme

This is a project which is mainly one of theoretical physics, and its objectives are in basic research; it does not have as for now any direct connection to industry. However, many of the themes developed are of enormous interest for several industrial activities: clustering of data for the biotechnologies and for financial applications, new error correcting codes for the computer and telecommunication industries, and models of financial markets for the banking industry. It is highly probably that some closer connections will be established with the industry during the course of the program. Furthermore, the main aim of the project – i.e. the training of physicists to the study of these newly emerging fields of science, will have an indirect effect in terms of cross fertilisation between academia and industry in the future.

For the post-docs working on the theme of 'Interacting agents in financial markets and socio-economy', there will be a possibility to spend a training period in the company 'Science & Finance', located in Paris and co-headed by J.P. Bouchaud, to complement their training by having access to large amounts of data and concrete problems from the financial industry.

13. Financial information

The proportion of the budget allocated to the temporary appointment of young researcher is 69 %, in agreement with the priority which we have given to the training of young post-docs.

Financial information on the network project				
Participant	Personnel and mobility costs related to the appointment of young researchers (euro) (A)	Costs linked to networking (euro) (B)	Overheads (euro) (C)	Totals (euro)
1.	129,600	24,000	31,400	185,000
2.	82,500	16,500	17,000	116,000
3.	80,208	12,592	23,200	116,000
4.	84,860	13,000	18,140	116,000
5.	81,000	15,667	19,333	116,000
6.	115,000	30,000	29,000	174,000
7.	122,400	22,600	29,000	174,000
8.	137,268	20,900	15,832	174,000
9.	108,000	37,000	29,000	174,000
10.	69,600	27,000	19,400	116,000
Totals	1,010,436	219,259	231,305	Grand Total 1,461,000