Hyperplane Annealing and Activator-Inhibitor-Systems

Thomas Laußermair

Technical University Munich and Siemens Nixdorf Information Systems AG, BU HLD FT 5,
D-8011 Poing, Siemensallee 2, Postfach 1260; Email: thomas@sun17.zfe.siemens.de

Abstract

This paper introduces a new optimization technique called hyperplane annealing. It is similar to the mean field annealing approach to combinatorial optimization. Both annealing techniques rely on a parallel relaxation dynamic. A connection is shown to a formal model of selforganized pattern formation, the activator-inhibitor-model. The unifying principle of all these relaxation models is a mechanism for the modulation of the non-linearity of the relaxation dynamic: Activator-inhibitor-systems show spatial modulation by diffusion, whereas the optimization approach uses functional modulation by gradients as well as temporal modulation by annealing of nonlinearity controlling system parameters. The new hyperplane annealing technique combines these modulation mechanisms within an algorithmic formulation with smaller computational complexity than mean field annealing. At a critical nonlinearity a phase transition leads to selforganized pattern formation in the relaxation matrix with the resulting structure corresponding to a solution of the optimization problem. As a concrete example the hyperplane annealing technique is used to solve instances of the NP-complete graph coloring problem.

Keywords

Combinatorial optimization, graph coloring problem, annealing principle, activator-inhibitor-model, relaxation dynamic, nonlinearity modulation, phase transition, hyperplane annealing.

1. INTRODUCTION: PROBLEM SOLVING PRINCIPLES GLEANED FROM NATURE

1.1. The relaxation principle - a physical metaphor

Physical systems can be described as parallel systems consisting of elementary particles with local forces determining their dynamic. An energetically closed system will tend to relax the local forces and approach an equilibrium state of minimal potential energy. This relaxation principle is exemplified by extremal surfaces of fluids or pneus.

From a mathematical point of view these states can be described as attractors of the relaxation dynamic. Relaxation into equilibrium states can be interpreted as iterative gradient descent: The local forces (gradients) vanish at local optima of the energy (potential function). Relaxation has been studied as a computational paradigm for solving large constraint satisfaction problems, in particular for cognitive tasks like vision.

> "Relaxation is the dominant mode of computation. ... the primary mode of computation in the brain is best understood as a kind of relaxation system .. in which the computation proceeds by iteratively seeking to satisfy a large number of weak constraints. .. The system should be thought of more as settling into a solution than calculating a solution." ([PDP86], p.135)

Two particular interesting tasks are the associative memory and discrete combinatorial optimization. Hopfield was among the first to introduce a discrete neuronal network formulation based on a spinglas model ([HOPF82]).

1.2. The annealing principle - a thermodynamic metaphor

Kirkpatrick et al. have described an optimization technique to improve gradient search methods called simulated annealing ([KIRK83]), which has been widely used and refined since then ([IJCNN89], [ICNC90], [PPSN90]). In the original thermodynamically inspired formulation a system parameter called temperature is iteratively reduced according to a socalled cooling schedule. This parameter determines the probability of accepting iteration steps in counter-gradient directions according to a Boltzmann distribution. Thus the temperature controlles the possibility of escaping local optima.

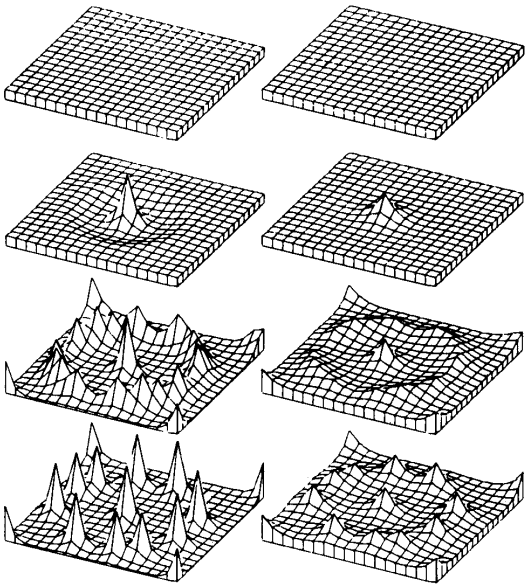
The authors of [BOUT89] describe a technique called mean field annealing which applies the annealing principle within the relaxation framework of a neuronal network. It can be described as a continuous generalization of the discrete Hopfield spinglas model integrating the annealing principle. It has been successfully applied to combinatorial optimization problems like graph partitioning ([PETE88]). The mean field annealing relaxation dynamic derives from a Boltzmann distribution controlled by a temperature-parameter. The hyperplane annealing relaxation dynamic derives from a population genetic formulation containing an exponent explicitly controlling the nonlinearity.

1.3. The activator-inhibitor-principle - a biochemical metaphor

The activator-inhibitor-system is a formal model of selforganized pattern formation proposed by Alan Turing in a study on the problem of biological morphogenesis. The model consists of two interacting and diffusing biochemical substances: An activator a and an inhibitor h. A system of two coupled nonlinear differential equations describes the temporal evolution of the two reactants. Here the form according to [MEIN82] will be given:

$$\dot{a} = \frac{a^2}{h} - \upsilon_a a + D_a \frac{\partial^2 a}{\partial x^2} + \rho_a \qquad (1a) \qquad\qquad \dot{h} = a^2 - \upsilon_h h + D_h \frac{\partial^2 h}{\partial x^2} + \rho_h \qquad (1b)$$

with a^2 auto- and cross-catalysis, $\frac{1}{h}$ cross-inhibiton, $\upsilon_{a/h}$, $D_{a/h}$ and $\rho_{a/h}$ the decay-rate, diffusion-rate and spontaneous production of activator / inhibitor.



activator inhibitor

Picture 1: Pattern formation after central activator induction ([HAKE83]).

The activator-inhibitor-model has been studied in detail by Meinhardt [MEIN82]. It explains a variety of biological pattern formation processes ranging from plants to insect embryos and sea shells. The activator-inhibitor-model describes a robust, parallel and fault-tolerant selforganized process of pattern formation. Picture 1 shows pattern formation after central activator induction (top-left) with subsequent temporal evolution from top to bottom (according to [HAKE83]). The patterns are not prespecified in detail but have characteristic properties depending on the parameters of the model and the coupling mechanism within the field or tissue. From a biochemic point of view this adds to the plausibility of the activator-inhibitor-system as a model of morphogenesis, since only several key parameters can be specified genetically.

From a mathematical point of view the growth processes during pattern formation depend primarily on the nonlinearity of the relaxation dynamic (1).

The properties of selforganized patterns can therefore be controlled to some extent by modulating the non-linearity of the dynamic. The selforganized patterns are unpredictable in detail due to the stochastic ρ-terms, but obey certain characteristics such as minimal and maximal mutual distance of activity centers. For example, the resulting activator pattern of the picture 1 can be regarded as satisfying a set of (spatial) constraints similar to the "k-out-of-N-constraint" discussed in [TAGL89]. The relaxation approach of the activator-inhibitor-model can be interpreted as a parallel procedure of generating such patterns satisfying a set of constraints. From this point of view one is tempted to use the relaxation principle for constraint optimization.
Obviously one has to convert the spatial constraints of the activator-inhibitor-model to the functional and logic constraints of the optimization problem. Furthermore the biochemical diffusion mechanism into neighbouring sites has to be replaced by an abstract coupling mechanism connecting different 'sites' of a suitable data structure. The following chapter describes the common framework to apply the relaxation principle to discrete optimization problems.

2. COMBINATORIAL OPTIMIZATION, DISTRIBUTION SPACE AND RELAXATION MATRIX

2.1. The graph coloring problem

The graph coloring problem (GCP) is a typical NP-complete constraint satisfaction problem ([GAJO79]). An instance of the GCP is given by a graph $G = (V,H)$ with finite vertex-set V ($|V|=N$), finite edge-set H and a positive number of M colors ($M \in \mathbb{N}^+$).
A mapping $B:V \rightarrow \{1,..,M\}^N$ is called a coloring of the graph. The set of all M^N different mappings constitutes the search space of all candidate colorings. A coloring B is permissible, if no two adjacent vertices of the graph G=(V,H) have the same color:

$$\text{Coloring B is } \textit{permissible} \quad \Leftrightarrow \quad (\ \{i,j\} \in H \text{ or } \{j,i\} \in H\ ) \ \Rightarrow \ B(V_i) \neq B(V_j)$$

A coloring is a solution of the GCP iff it is permissible.

2.2. The relaxation approach

The search space can be transformed to a space of distributions among all possible colorings. Let the variables V_i denote the vertices of the graph and the values W_k correspond to the colors of the vertices. Let the value $x_{ik} \in [0,1]$ denote the *activity* of value W_k at variable V_i. The activities x_{ik} constitute a relaxation matrix X. Usually the activities of a single variable, i.e. the rows of the matrix satisfy the *simplex-condition*: $\sum_{k=1}^{M} x_{ik} = 1$.
Every coloring B corresponds to a *row-degenerate distribution* X within the relaxation matrix:

$$\text{Distribution X is } \textit{row-degenerate} \quad \Leftrightarrow \quad \forall i: \ \exists k_i: \ x_{ik_i}=1 \ \text{ und } \ x_{il}=0 \text{ für } l \neq k_i.$$

The central element of the relaxation approach is a relaxation dynamic defined on the relaxation matrix X. The basic idea of using the relaxation approach for optimization is to start with a uniform activity distribution with some noise added: $x_{ik}=M^{-1}+\delta$. The relaxation dynamic will upon iteration lead to stable attractors. A suitable dynamic will have row-degenerate distributions as stable attractors corresponding to colorings.
Solving a combinatorial optimization problem in the distribution space concept thus amounts to pattern formation in the relaxation matrix from uniform initial distributions to structured row-degenerate final distributions. For the optimization problem further restrictions have to be imposed upon the relaxation dynamic: The stable attractors should be functionally optimal and/or satisfy all constraints.
Satisfying a set of constraints can be converted into minimizing a function of collective constraint violation. Therefore the common approach for solving constraint optimization problems with the relaxation approach is a two-step process:

First, a suitable potential or energy function is designed with minima at permissible states. For a large class of problems including GCP the energy-function can be formulated as sum of all pairwise activity couplings. The following scheme represents these activity couplings as product of the coupling factors c_{kl} between values W_k and W_l and d_{ij} between variables V_i and V_j:

$$E(X) = \sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij} \sum_{k=1}^{M}\sum_{l=1}^{M} c_{kl}\, x_{ik}\, x_{jl} \qquad (2)$$

Second, a suitable relaxation dynamic is designed, usually containing partial derivatives of the potential function to achieve a gradient descent ([HOPF82], [HOTA85]). Hopfield characterizes the differential equation of the relaxation dynamic as a program for solving the optimization problem.

2.3. Formulation of the graph coloring problem within the relaxation approach

A problem instance will be described by its potential or energy function. According to (2) this amounts to specifying the coupling factors d_{ij} and c_{kl}. Within the relaxation framework an instance of the GCP is given by a quadrupel (V,W,C,D):

V: $V=\{V_i\}$, $i \in \{1,..,\mathbf{N}\}$; V is a finite set of *variables*;

W: $W=\{W_k\}=\{1,2,...,M\}$, $k \in \{1,..,\mathbf{M}\}$; W is a finite set of discrete *values*;

C: $C = (c_{11}, .. , c_{kl} , .. , c_{MM})$, $k,l \in \{1,..,M\}$,

C is a M*M-matrix of *value coupling factors* $c_{kl} = \begin{cases} +1 & \text{if } l=k \\ 0 & \text{otherwise} \end{cases}$

D: $D = (d_{11}, .. , d_{ij} , .. , d_{NN})$, $i,j \in \{1,..,N\}$

D is a N*N-matrix of *variable coupling factors* $d_{ij} = \begin{cases} +1 & \text{if } \{i;j\} \in H \\ 0 & \text{otherwise} \end{cases}$

For the GCP only those variable coupling factors d_{ij} are non-zero for which $\{i,j\} \in H$ (vertices V_i and V_j are connected within the graph G). Furthermore only those value coupling factors c_{kl} are non-zero which represent the same color k=l. Thus equation (2) reduces to

$$E_{GCP} = \sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij} \sum_{k=1}^{M} x_{ik}\, x_{jk} = \sum_{\{i;j\} \in H} \sum_{k=1}^{M} x_{ik}\, x_{jk} \qquad (3)$$

However, the energy function has to incorporate not only functional but also structural terms. A common scheme uses weighted cost and structure terms.

$$E = \alpha\, E_{structure} + \beta\, E_{cost} \qquad (4)$$

One of the problems with this approach is to balance the relative influence of structure and cost. The authors of [PETE88] showed that robust and general formulations for relaxation into permissible (structure) and near-optimal (cost) solutions cannot be achieved by merely tuning the relativ weights α and β for structural and functional constraints. For the experiments in this paper we will set the weights α and β both to 1 and use the common structure term

$$E_{structure} = \sum_{i}\Big[\sum_{k} x_{ik} - 1\Big]^2 \text{ (rows)} + \sum_{k}\Big[\sum_{i} x_{ik} - 1\Big]^2 \text{ (columns)} \qquad (5)$$

3. RELAXATION DYNAMIC AND MODULATION OF ITS NONLINEARITY

3.1. A relaxation dynamic for optimization

Voigt has proposed the following population genetic oriented formulation of a relaxation dynamic ([VOIG89]):

$$x_{ik}' = x_{ik}\frac{f_{ik}(X)}{\Phi_i} \quad \text{with} \quad f_{ik} = \frac{\partial E}{\partial x_{ik}} \quad \text{and} \quad \Phi_i = \sum_{k=1}^{M} x_{ik} f_{ik} \qquad (6)$$

The activity x_{ik} and the term $f_{ik}(X)$ are interpreted as the relative frequency and the fitness of allele W_k for gene V_i within a population X. Equation (6) describes a selection process which can lead to a 'winner-takes-all' selection of single alleles for every gene, corresponding to row-degenerate distributions. The term Φ_i is used to normalize the relative frequencies (simplex-condition) and is interpreted as the average fitness of gene V_i.

3.2. Connection to the activator-inhibitor-model

Voigt observed the stability of uniform distributions as well as new stable attractors of equation (6) not corresponding to row-degenerate distributions. Therefore he introduced a parameter called selection pressure u to modulate the nonlinearity of (6) ([VOIG89]):

$$f_{ik}' = \frac{\partial E}{\partial x_{ik}} + u\,x_{ik} = f_{ik} + u\,x_{ik} \qquad (7)$$

Using equation (6) in its differential form we obtain

$$\dot{x}_{ik} = x_{ik}\frac{(f_{ik}' - \Phi_i)}{\Phi_i} = x_{ik}\frac{(f_{ik} + ux_{ik} - \Phi_i)}{\Phi_i} = \frac{u}{\Phi_i}x_{ik}^2 + \frac{(f_{ik} - \Phi_i)}{\Phi_i}x_{ik} \qquad (8)$$

which compares to equation (1a) of the activator-inhibitor-model. In both cases a quadratic term can lead to a winner-takes-all selection depending on a modulating factor of the quadratic term.
The normalization term Φ_i is analogue to the cross-inhibition effect of the inhibitor h in (1a). Neglecting spontaneous production ρ and decay rates υ, the two models are formally identical under the limit $D_a \to 0$ and $D_h \to \infty$: Under infinite diffusion-rate the quantity h shows no spatial structure and is equal to the normalization average.

3.3. Modulation of the nonlinearity of the relaxation dynamic

The process of pattern formation in relaxation models depends upon the nonlinearity of the relaxation dynamic. As a unifying principle, three different mechanisms for modulation of this nonlinearity can be identified:

Spatial modulation by diffusion: In the activator-inhibitor-model the exponent of the quadratic term in (1a) is implicitely modulated by spatial diffusion and the effect of cross-inhibition. Suppose that neither activator nor inhibitor are diffusing. Then, since h is cross-catalyzed proportional to a^2 in (1b), the quadratic term of (1a) degenerates to a constant term and the negative feedback of the second decay term dominates. Now if the substrate is sufficiently large and the inhibitor h diffuses more quickly than the activator ($D_h > D_a$), then a local excess of activator can lead to auto-catalysis and positive feedback: In this case the inhibitor h is locally smaller than a^2 due to its rapid diffusion into neighbouring sites. The first term of (1a) will therefore be equal to a value of a^p with spatial varying exponent p. This quantity may locally exceed the effect of the decay term and allow for autocatalytic growth of an activator center.

Functional modulation by gradients: The term $f_{ik} = \frac{\partial E}{\partial x_{ik}}$ in (6) introduces gradient information into the relaxation dynamic leading to positive feedback for activities with fitness above average (selection). Functional modulation by gradients is important, but leads to the problem of local optima due to the couplings between different activities and the nonlinearity of the energy function E. This problem is quite fundamental and cannot be eliminated by tuning the weight parameters α and β in equation (4) alone (see comments above).

Temporal modulation by annealing: As an additional technique to overcome local optima most approaches incorporate the principle of simulated annealing ([KIRK83]) as briefly introduced above.
Voigt's approach (8) uses explicit modulation of a factor for a term with fixed degree of nonlinearity (quadratic term). Increasing this "selection pressure" factor according to an annealing schedule leads to stable attractors close to row-degenerate distributions. However, there are some problems with stable semi-degenerate distributions, i.e. two alleles co-stabilize at x_{ik}=0.5. This phenomenon has been experimentally studied in more detail in [LAUS92]. A connection to Eigen's hypercycles is conjectured ([EIGE71]).
Hyperplane annealing as introduced below uses explicit modulation of the exponent of a term with variable degree of nonlinearity. This direct control of the nonlinearity makes it even more similar to the activator-inhibitor-model.

4. HYPERPLANE-ANNEALING

4.1. Nonlinear damping on curved hyperplanes: Generalizing the simplex condition

In the approaches encountered sofar the relaxation dynamic contains normalization terms to enforce the simplex-condition. Geometrically speaking, the state space of the distribution X_i of every optimization variable (i.e. row of the relaxation matrix) corresponds to the M-dimensional unit simplex. The planar geometry of the simplex implies a linear damping characteristic of competing values W_k for a variable V_i: An increase of a particular activity x_{ik} leads to a proportional linear decrease of competing activities x_{ij} ($j \neq k$). The state space of the distribution X_i can be generalized to a curved hyperplane. With a reversible function g(x) we extend the simplex condition to the following hyperplane condition: $\sum_{j=1}^{M} g(x_{ij}) = 1$. Equation (6) then changes to

$$x_{ik}' = \frac{x_{ik} f_{ik}}{g^{-1}\left[\sum_j g(x_{ij} f_{ij})\right]} \qquad (9)$$

where the hyperplane condition is easily verified. Using the simple polynom $\boxed{g(x) = x^k}$, $k \in \mathbb{R}^+$, we get curved hyperplanes with k as the curvature exponent. At k = 2 the hyperplane coincides with the surface of the unit sphere ($x_i \in [0,1]$).
In [LAUS92] the nonlinear damping characteristic on curved hyperplanes with k>1 is experimentally and formally verified. The intuitive motivation for this is the following: In the vicinity of the uniform distribution, i.e. at the center of gravity of the hyperplane, there is a weak selection tendency. In the vicinity of degenerate distributions, i.e. at the boundary of the hyperplane, the selection tendency is strong. The decision process is implicitely divided from weak at the beginning to strong at the end quite similar to the explicit enforcement of this characteristic by annealing. A weak initial selection tendency implicitely allows more decision alternatives to be regarded in the early relaxation phase.

4.2. Explicit modulation of nonlinearity

The above considerations of pattern formation in the activator-inhibitor-model suggest a nonlinear selection characteristic, which can perfectly complement the nonlinear damping characteristic of curved hyperplanes. A nonlinear selection characteristic can be achieved by introduction of an exponent p explicitely controlling the nonlinearity of the relaxation dynamic.

$$x_{ik}' = \frac{(x_{ik}f_{ik})^p}{\sum_j (x_{ij}f_{ij})^p} \qquad (10)$$

Combining (9) and (11) leads to

$$x_{ik}' = \frac{(x_{ik}f_{ik})^p}{g^{-1}\left[\sum_j g\{(x_{ij}f_{ij})^p\}\right]} \qquad (11)$$

Similar to the question about the relative size of the weight parameters α and β in (4), now arises the question about the proportion of p and k in (11). A natural answer is a reciprocal setting $p = \frac{1}{k}$, which also drastically simplifies the relaxation dynamic to

$$x_{ik}' = g^{-1}\left[\frac{x_{ik}f_{ik}}{\sum_j x_{ij}f_{ij}}\right] \qquad (12)$$

When compared to (6), the resulting form (12) differs only in the additional inverse function g^{-1}. It is important to note that the computational complexity of the relaxation dynamic is thus not increased.

4.3. Integration into the annealing framework

The above relaxation dynamic is now integrated into the annealing framework forming the new hyperplane annealing algorithm (see below). Its computational runtime complexity has been investigated in detail in [LAUS92]. Since the algorithm uses explicit gradient information (in step 2.1.a), two observations can be made: The computational complexity

- is in general less than that of mean field annealing, which has to discover gradients and
- depends on the specific problem via the computational complexity of the gradients.

For a parallel implementation the computational complexity O(Algo) of hyperplane annealing (HPA) and mean field annealing (MFA) are summarized in table 1 for the traveling salesman problem (TSP) and graph coloring problem (GCP).

Table 1: Computational runtime complexity of mean field annealing (MFA) and hyperplane annealing (HPA) for TSP und GCP.

O(Algo)	$MFA_{parallel}$	$HPA_{parallel}$
TSP	$O(N^4)$	$O(N^2)$
GCP	$O(M*N^2)$	$O(M*N)$

Hyperplane-Annealing

1. Initialize curvature k of hyperplane, exponent $p = \frac{1}{k}$ of relaxation dynamic and N*M-relaxation matrix X

with $x_{ik} := (\frac{1}{M})^p + \delta$

2. Repeat until distribution is row-degenerate:

2.1. Repeat until equilibrium:

a. Compute fitnesses $f_{ik} := \frac{\partial E}{\partial x_{ik}}$ b. Compute activities $x_{ik} := x_{ik} f_{ik}$

c. Normalize rows: $x_{ik} := \frac{x_{ik}}{\sum_j x_{ij}}$ d. Project to hyperplane: $x_{ik} := g^{-1}(x_{ik}) = (x_{ik})^p$

2.2. Reduce the curvature k. Increase the autocatalysis exponent $p = \frac{1}{k}$.

3. Transform row-degenerate distribution into a solution.

δ is small random noise. For Steps 2. and 2.1. some termination criterium has to be defined, usually in terms of the maximal change Δx_{ik} being less than a small constant ε.

5. EXPERIMENTS

5.1. Coloring planar graphs and maps

The hyperplane annealing technique will be applied to solve instances of the graph coloring problem. The graphs chosen are special planar graphs. The problem can be best visualized using dual maps (graph-vertices = map-countries, graph-edges = map-borders). A permissible colored graph corresponds to a colored map with no two bordering countries having the same color. The maps chosen are periodic maps consisting of regular hexagons and aperiodic maps built out of two types of penrose tiles, which allow only aperiodic tilings of the 2-dimensional plane. The degree of connectivity of the corresponding graphs is six and four, resp.; the required minimal degree for the graph coloring problem to be NP-complete is four ([GAJO79]). The concrete instances used for experimentation consist of 64 hexagons and 78 penrose tiles.

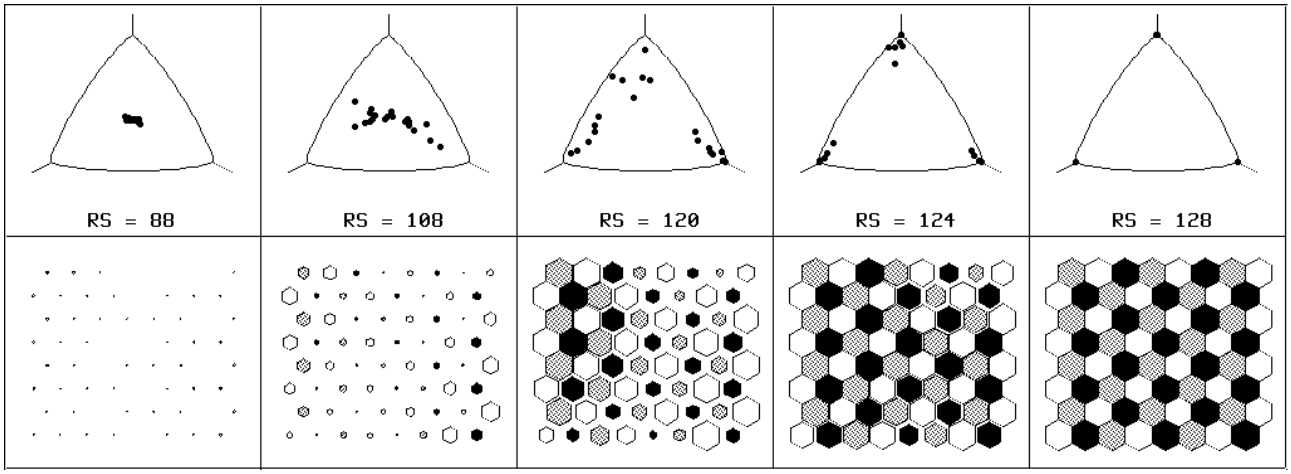


Figure 2: Phase transition in hyperplane annealing applied to the 64-hexagon graph coloring problem.

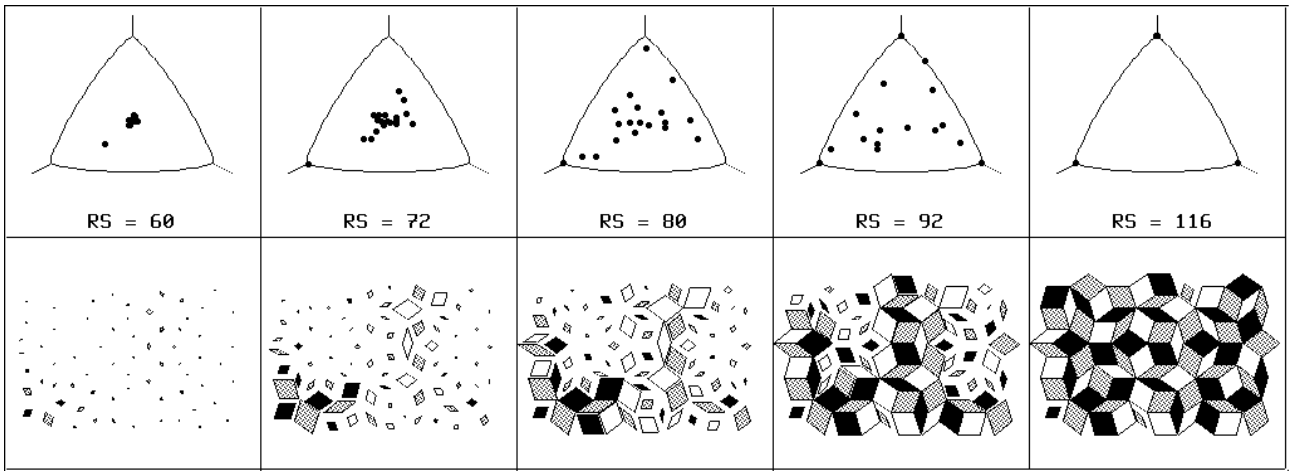


Figure 3: Phase transition in hyperplane annealing applied to the 78-penrose graph coloring problem.

The row-distributions within the relaxation matrix can be visualized as a point on the curved hyperplane. The movement of the 64 (78) points is shown in the top half of figure 2 (3). RS is the number of iterations of the relaxation dynamic within step 2.1. Step 2. is iterated only once with the experimentally determined critical curvature k=1.3. The bottom half indicates the colored map; the size of the countries indicates the amount of structure in the resp. row, i.e. the distance of the resp. point from the center of gravity of the hyperplane. Finally all points are attracted towards one corner of the hyperplane: The row-degenerate distribution assigns one color to the corresponding country.

5.2. Phase transitions and frustrations

Both hyperplane and mean field annealing show a critical value of curvature and temperature: There the initial uniform distribution becomes unstable and during a single relaxation (step 2.1) a sudden structure formation within the relaxation matrix takes place. For some optimization problems this suffices to produce a matrix structure close to a row-degenerate distribution. In [BOUT89] the problem of bi-partitioning graphs is thus solved without annealing. Similarly, the GCP-instances (N=64, M=3) colors can be solved without annealing.
The mathematical reason for this is that the stability of the attractors of the relaxation dynamic changes at the critical nonlinearity. The annealing process starts with high curvatures (HPA) and high temperatures (MFA). In the beginning the uniform distributions $x_{ik}=M^{-p}$ (HPA) and $x_{ik}=M^{-1}$ (MFA) are stable attractors. Below the critical values of curvature and temperature this attractor becomes unstable and with further decreasing values row-degenerate distributions become stable attractors. This process can be visualized by plotting the distributions of each variable X_i as a point on the simplex (MFA) or curved hyperplane (HPA). The relaxation at or below the critical value of the system parameters temperature and curvature leads to a trajectory of these points beginning at the center of gravity of the hyperplane and ending in one of its corners, each corresponding to one of the possible colors (see figures 2 and 3). By stability analysis of the relaxation dynamic a critical temperature ([BOUT89]) or curvature ([LAUS92]) can be formally derived, at which a phase transition leads to row-degenerate distributions without annealing, so that the runtime complexity can be drastically reduced.
A permissible coloring corresponds to a global optimum in the search space of all distributions. Local optima correspond to colorings with some frustrations, i.e. minor violations of the coloring rule. Due to the local optimum problem the likelihood of frustrations increases strongly with growing graph sizes for both annealing methods (MFA and HPA). Whereas in the case of the 64-hexagon-map frustrations occur sometimes, the colorings for a 256-hexagon-map always showed frustrations typically separating four areas of locally permissible colorings. This phenomenon has been investigated in more detail in [LAUS92]; first results indicate a superior scaling behavior for hyperplane annealing as compared to mean field annealing.

6. CONCLUSION

Hyperplane annealing is a new technique for combinatorial optimization similar to mean field annealing. It uses a different form of relaxation dynamic that is inspired by the activator-inhibitor-model of selforganized pattern formation. The main results are twofold: Hyperplane annealing uses gradients and has therefore less computational complexity than mean field annealing. The solution quality of both techniques is very similar with hyperplane annealing showing a better scaling behavior. Further work on these annealing techniques should incorporate hierarchical relaxation in a multilayer architecture to avoid the problems of frustration. A promising step in this direction has been proposed by [VOIG90].

7. BIBLIOGRAPHY

[BOUT89] D.van den Bout, T.Miller
"Graph Partitioning using Annealed Neural Networks", in: IJCNN 1989, I, 521-528

[EIGE71] M.Eigen, "Selforganization of Matter and the Evolution of Biological Macromolecules"
Die Naturwissenschaften, Heft 10 Oktober, S.465-523, 58.Jahrgang, 1971

[GAJO79] M.Garey, D.Johnson
"Computers and Intractability", A Guide to the Theory of NP-Completeness
W.H.Freeman and Company, San Francisco 1979

[HAKE83] H.Haken, "Synergetik", Springer Verlag, Berlin 1983

[HOPF82] J.Hopfield
"Neural networks and physical systems with emergent collective computational abilities"
Proc. of the National Academy of Sciences, 79:2554-2558, 1982

[HOTA85] J.Hopfield, D.Tank
"Neural computation of decisions in optimization problems"
Biological Cybernetics, Vol.52, S.141-152, Springer-Verlag 1985

[ICNC90] R.Eckmiller, G.Hartmann, G.Hauske (Eds)
"Parallel Processing in Neural Systems and Computers", Elsevier Science Publ., Amsterdam 1990

[IJCNN89] International Joint Conference on Neural Networks, IEEE TAB Neural Network Comitee,
Washington 18.-22.6.1989

[KIRK83] S.Kirkpatrick, C.D.Gelatt, M.P.Vecchi
"Optimization by simulated annealing", Science, 220, pp.671-680, 1983

[LAUS92] T.Laussermair, "Hyperflächen-Annealing. Ein paralleles Optimierungsverfahren basierend auf
selbstorganisierter Musterbildung durch Relaxation auf gekrümmten Hyperflächen"
Doctoral Dissertation, Technical University Munich, 1992

[MEIN82] H.Meinhardt, "Models of biological pattern formation", Acadenic Press, London 1982

[PDP86] D.Rumelhart, J.McClelland,(Eds)
"Parallel Distributed Processing", Volumes I and II, MIT Press 1986

[PETE88] C.Peterson, J.Anderson, "Neural Networks and NP-complete Optimization Problems;
A Performance Study on the Graph Bisection Problem", Complex Systems 2 (1988) 59-89

[PPSN90] R.Männer, H.P.Schwefel (Eds)
"Parallel Problem Solving from Nature", Conference Proceedings, Springer Verlag, 1990

[TAGL89] G.A.Tagliarini, E.W.Page
"Learning in systematically designed networks", in: IJCNN 1989, I, 497-502

[VOIG89] H.M.Voigt, "Optimization by Selection Pressure Controlled Replicator Networks"
System Analysis, Modelling & Simulation 6 (1989) 4, 267-278

[VOIG90] H.M.Voigt, I.Santibanez-Koref, "Multi-Layer Replicator Network Algorithms"
Evolution and Optimization '89, pp.95-106, Akademie-Verlag Berlin, 1990