Proc. Int. Conf. on Artificial Neural Networks ICANN 1999, Edinburgh, pages 868-873

BACKTRACKING DETERMINISTIC ANNEALING FOR CONSTRAINT SATISFACTION PROBLEMS

Heiko Wersing and Helge Ritter

University of Bielefeld, Faculty of Technology
P.O.-Box 10 01 31, D-33501 Bielefeld, Germany
{hwersing,helge}@techfak.uni-bielefeld.de

Abstract

We present a new deterministic annealing approach to the solution of quadratic constraint satisfaction problems with complex interlocking constraints, such as exemplified in polyomino tiling puzzles. We first analyze the dynamical properties of the solution strategies implemented by deterministic annealing (DA) in the analog neural representation of Potts-Mean-Field (PMF) and penalty-function-based competitive layer model (CLM) neural networks, revealing a similar mechanism. The key idea of our extension of these plain DA approaches is motivated by classical backtracking algorithms. We show that their ability for iterative local pruning of the search space can be implemented within the framework of DA by introducing local temperature parameters which are "reheated" when locally unresolved conflicts occur. To achieve the pruning of the search space, reheating is accompanied by a modification of the constraint-implementing weight matrix to reduce the chance of reentering the same local configuration. The weight changes provide a learning mechanism that facilitates the generation of a solution for subsequent runs. We demonstrate the benefits of the resulting "backtracking deterministic annealing" algorithm (BDA) by applying it to a pentomino tiling problem. We show that the method reliably finds perfect solutions to the task, while the plain DA approach for both PMF and CLM is unable to solve the task in a comparable or even considerably larger number of iterations.

1 Introduction

Combinatorial optimisation problems have a long tradition in the neural network field dating back to the seminal work of Hopfield and Tank [4]. While the early approaches suffered from problems of spurious and local minima, the later developed deterministic [8] or Potts-Mean-Field annealing networks [6] provide techniques which are competitive with algorithms like simulated annealing with respect to computation time and solution quality. However, the complexity of the constraints can pose problems requiring the development of special techniques. An example is the matching (or "two-way") assignment constraint in the Travelling Salesman problem [5]. Polyomino tiling problems [3] provide much more complex interlocking constraints that challenge existing DA approaches even further, since they may require complex reconfigurations, when local constraints become violated in later stages of the annealing phase.

Backtracking has been very successfully applied to tasks of this type [1]. This suggests that it might be possible to find an algorithm that can combine some of the benefits of backtracking with the superb capabilities of DA algorithms to explore complex search spaces. In the present contribution, we suggest an approach that extends DA by introducing a local temperature control scheme which is accompanied by modifications of the constraint-implementing weight matrix to reduce the chance of reentering the same local configuration.

The plan of the paper is as follows. Section 2 gives a brief overview over two deterministic annealing approaches and discusses their dynamical search strategy. In Section 3 we introduce the "backtracking deterministic annealing" (BDA) which we apply to the problem of pentomino tilings in Section 4 and discuss our results in Section 5.

2 Constrained Optimisation with Neural Networks

We consider constrained optimisation problems which can be stated by the minimisation of a quadratic energy

$$E = -\frac{1}{2} \sum_{rr'\alpha\beta} \mathsf{f}^{\alpha\beta}_{rr'} x_{r\alpha} x_{r'\beta} \rightarrow \text{min.} \qquad (1)$$

subject to the weighted labeling constraint

$$\sum_{\alpha} x_{r\alpha} = 1, \qquad (2)$$

and $x_{r\alpha} \in \{0, 1\}$. $x_{r\alpha}$ denotes the certainty of assigning a label $\alpha \in \{1, \ldots, L\}$ to a feature r

among L alternatives and $\mathsf{f}_{rr'}^{\alpha\beta}$ encodes the mutual compatibility of assigning label α to feature r and label β to feature r'. We denote the L neurons $x_{r\alpha}$, $\alpha \in \{1, \ldots, L\}$ as the *column* r.

The classical Potts-Mean-Field neural network approach to the solution of this problem [6] is given by minimising the mean-field free energy F_{PMF} with

$$F_{\text{PMF}} = E + T \sum_{r\alpha} x_{r\alpha} \log x_{r\alpha} \qquad (3)$$

subject to (2). If we interpret $x_{r\alpha}$ as the probability of labeling r with α, then the additional entropy term constrains the randomness of the weighted labeling. The actual solutions are obtained by iterating the following dynamics (parallel, serially, or asynchronously [6]) together with a gradual lowering of the temperature T

$$\dot{x}_{r\alpha} = -x_{r\alpha} + \frac{\exp(-TE_{r\alpha})}{\sum_\beta \exp(-TE_{r\beta})}, \qquad (4)$$

where

$$E_{r\alpha} = -\frac{\partial E}{\partial x_{r\alpha}} = \sum_{r'\beta} \mathsf{f}_{rr'}^{\alpha\beta} x_{r'\beta} \qquad (5)$$

is the negative local gradient of the original energy. For fixed T the dynamics (4) always converges to a local minimum of the free energy (3). For high T the free energy is dominated by the convex entropy term with unique minimum at $x_{r\alpha} = 1/L$. Lowering T, the free energy gains more and more of its original structure until for $T \to 0$ the original energy E is recovered.

The Competitive Layer Model (CLM) [7, 11] considers a different energy of the form

$$F_{\text{CLM}} = E + J \sum_r \Big(1 - \sum_\alpha x_{r\alpha}\Big)^2 + T \sum_{r\alpha} x_{r\alpha}^2. \qquad (6)$$

The weighted labeling constraint (2) is enforced by an additional penalty function. The logarithmic entropic constraint term is replaced by a quadratic diagonal term. In fact, any convex function of $x_{r\alpha}$ could achieve the same mechanism of biasing the minima of F towards the uniform labeling $x_{r\alpha} = 1/L$. The quadratic choice in (6) leads to a piecewise linear gradient descent dynamics of the form

$$\dot{x}_{r\alpha} = J\Big(1 - \sum_\beta x_{r\beta}\Big) + E_{r\alpha} - T\, x_{r\alpha} \;, \qquad (7)$$

subject to the simple non-linearity $x_{r\alpha} \geq 0$. The importance of the diagonal self-interacting terms was already stressed in [6], however, only exploited in the final stages of annealing where $T \to 0$. In a slight generalisation of previous results [11, 10], the following theorem states conditions which exclude spurious attractors of the CLM with respect to the labeling task

Theorem 1. If the compatibilities satisfy $\mathsf{f}_{rr}^{\alpha\alpha} + \mathsf{f}_{rr}^{\beta\beta} > 2\,T + 2\mathsf{f}_{rr}^{\alpha\beta}$ for all r, α, β, and $J > \max_{r\alpha} \sum_{r'\beta} \max(0, \mathsf{f}_{rr'}^{\alpha\beta})$, then all local minima of F_{CLM} are unambiguous labelings with i) at most one positive activity $x_{r\alpha(r)} = 1 + E_{r\alpha(r)}/J$ in a column, where $\alpha(r)$ is the index of the maximally supporting label with $E_{r\alpha(r)} > E_{r\beta\neq\alpha(r)}$ or ii) for all labels in a column there holds $x_{r\alpha} = 0,\ E_{r\alpha} \leq 0$.

The theorem states that the dynamics (7) converges towards an unambiguous labeling. Due to the "soft" enforcing of the weighted labeling constraint, however, the final output $x_{r\alpha(r)}$ is modulated by an amount given by $E_{r\alpha(r)}/J$. While this can be rather usefully exploited in certain vision labeling tasks [10], here we choose J sufficiently large such that the deviations are small and the final output state can be simply rectified by $x'_{r\alpha} = \theta(x_{r\alpha})$ where $\theta(0) = 0$ and $\theta(x) = 1$ for $x > 0$.

We now discuss the dynamical properties of the two DA algorithms for PMF and CLM. We show that a comparison of the linearised dynamics reveals a similar structure at the dynamical level though using different means of enforcing constraints at the energy function level.

In the limit of large T both energies (3) and (6) share a common global minimum $x_{r\alpha} = 1/L$ which is then the only attracting fixed point of both dynamics (4) and (7). Linearising the PMF dynamics about this fixed point gives

$$\dot{x}_{r\alpha} = \frac{1}{L} - x_{r\alpha} + \frac{1}{LT}\Big(E_{r\alpha} - \frac{1}{L}\sum_\beta E_{r\beta}\Big). \qquad (8)$$

To compare the two vector fields (7) and (8), a projection onto two orthogonal subspaces is helpful where we consider the DC subspace with $\dot{x}_{r\alpha}^{\text{DC}} = 1/L \sum_\beta \dot{x}_{r\beta}$ and the AC subspace with $\dot{x}_{r\alpha}^{\text{AC}} = \dot{x}_{r\alpha} - \dot{x}_{r\alpha}^{\text{DC}}$ respectively. While the AC subspace captures components which cause differences between labels α in a column r, the DC subspace carries the dynamics of the summed activity in a column r.

Instead of moving into the direct gradient direction $E_{r\alpha}$ the (entropic) constraints lead to the following effective directions

$$E_{r\alpha}^{\text{PMF}} = E_{r\alpha}/LT - x_{r\alpha}, \qquad (9)$$
$$E_{r\alpha}^{\text{CLM}} = E_{r\alpha} - T x_{r\alpha}.$$

With this notation of the components we get for PMF:

$$\dot{x}_{r\alpha}^{\mathrm{AC}} = E_{r\alpha}^{\mathrm{PMF}} - \frac{1}{L}\sum_{\beta} E_{r\beta}^{\mathrm{PMF}}, \tag{10}$$

$$\dot{x}_{r\alpha}^{\mathrm{DC}} = \frac{1}{L} - \frac{1}{L}\sum_{\beta} x_{r\beta}. \tag{11}$$

For the CLM we obtain:

$$\dot{x}_{r\alpha}^{\mathrm{AC}} = E_{r\alpha}^{\mathrm{CLM}} - \frac{1}{L}\sum_{\beta} E_{r\beta}^{\mathrm{CLM}}, \tag{12}$$

$$\dot{x}_{r\alpha}^{\mathrm{DC}} = J\Big(\frac{1}{L} - \frac{1}{L}\sum_{\beta} x_{r\beta}\Big) + \frac{1}{L}\sum_{\beta} E_{r\beta}^{\mathrm{CLM}}. \tag{13}$$

This shows that in the DC space both linear dynamics (for sufficiently large J) are driven towards the affine constraint surface (2). In the AC-subspace both dynamics have a very similar structure, differing mainly in the way the temperature T enters the dynamics in (9): For the PMF the effective matrix driving the AC dynamics is divided by LT with a constant negative identity matrix, while for the CLM the original gradient is kept and only the diagonal is shifted towards negative eigenvalues. The critical temperatures at which the fixed point gets unstable are given by

$$T_c^{\mathrm{PMF}} = \lambda_{\max}\{\mathsf{f}_{rr'}^{\alpha\beta} - 1/L\sum_{\beta}\mathsf{f}_{rr'}^{\alpha\beta}\}/L, \tag{14}$$

$$T_c^{\mathrm{CLM}} = \lambda_{\max}\{\mathsf{f}_{rr'}^{\alpha\beta} - J\delta_{rr'}\}. \tag{15}$$

If T falls below the critical T_c only the dominant eigenmodes in the AC subspace are expressed by the dynamics while other eigenmodes are suppressed due to the negative eigenvalue shift for both PMF and CLM. This is a central mechanism of the deterministic annealing approach as has also been discussed for clustering applications [8]. Therefore the success of the DA heavily depends on the property, that the maximum eigenvalue directions are directions within the search space with a higher probability of a global minimum occurring. We may consider this as the extraction of some global properties encoded implicitly into the compatibility weights of the optimisation problem.

For lower T the dynamics gets more complex and there are phases of linear growth within restricted configuration subsets at bifurcations and phases of tracking the local equilibria of the free energy which move continuously with respect to T towards the boundary. Due to the softmax nonlinearity the PMF performs a more smooth sliding of equilibria to the minima of E in the $T \to 0$ limit, while the "hard' nonlinearity introduced by the condition $x_{r\alpha} \geq 0$ causes a sharpened exploration of the "corners" of the state space. As is discussed in Section 4 this causes a more deterministic behaviour of the CLM algorithm for slow annealing.

3 Combining Backtracking and Deterministic Annealing

Backtracking algorithms [1] solve constraint satisfaction problems by iterative reduction of the combinatorial search tree whenever local constraint conflicts are detected. The effective pruning of the search space leads to a speed gain which makes them applicable to a wide range of constraint satisfaction problems [1]. While the local search for feasible solutions is exhaustive, however, global properties which follow from the structure of the problem, i. e. implicitly encoded into the weight matrix of the energy (1), are not directly exploited and can only be taken into account by hand by choosing an appropriate tree-search strategy. On the other hand, the success of deterministic annealing neural networks is based on extracting global information by dynamically exploring maximum eigenvalue directions of the search space. They can also be simulated on parallel and analog hardware, offering scalable speed gains. This motivates the formulation of a hybrid approach which combines the advantages of the two approaches.

We suggest the implementation of a backtracking deterministic annealing (BDA) network in the following way: The starting point is one of the DA approaches like PMF or CLM. The global temperature T is replaced by individual temperatures T_r for each feature r to allow for a locally controlled rearrangement of labels. In the beginning, annealing is performed for all T_r synchronously, however, when $T_r < T^{(1)}$, where $T^{(1)} \ll T_c$ is a typical temperature causing a pronounced decision for one of the labels, say (r, α), it is checked whether a local constraint conflict of (r, α) with another assigned (r', β) has occurred. If this is the case, the weight $\mathsf{f}_{rr'}^{\alpha\beta}$ is changed to $\mathsf{f}_{rr'}^{\alpha\beta} - \Delta$, causing an energy penalty for this subconfiguration. The local temperatures T_r and $T_{r'}$ are raised to an intermediate value $T^{(2)}$ which must be sufficiently high to restore a superimposed state for features r and r'. The chance of reentering the previously encountered erroneous state is now diminished due to the weight change. Performing these steps synchronously at multiple sites leads to a search procedure, which can effectively escape from local minima. The system "learns" to optimise its own weights for obtaining a solution to the optimisation problem.

Quite generally, constraint satisfaction problems like tiling problems are modelled by inhibitory interactions of conflicting alternatives. In that case the energy level of a proper solution forming a global minimum does not depend on the constraint-implementing weights and by changing these weight components only local minima are affected. The task of the BDA approach is not to optimise these weights with respect to the original problem alone but with respect to the problem *and* the search strategy implemented by the annealing dynamics.

We now state the explicit algorithm which we used for the simulation results of the BDA with the CLM as shown in the next section. The CLM dynamics (7) can be simulated in principle by standard differential equation integrators like the Euler or Runge-Kutta method and can be computed in parallel. The piecewise linear dynamics, however allows also for a sequential asynchronous update [10] which shows rapid convergence to the proper equilibria of (7) and can be very easily implemented:

1. Set $T_r(0) = T_c$ for all r, where $T_c = \lambda_{max}(\mathsf{f}^{\alpha\beta}_{rr'} - J\delta_{rr'})$. Initialise all $x_{r\alpha}$ with random values $x_{r\alpha}(t=0) \in [1/L - \epsilon, 1/L + \epsilon]$.

2. Do $N \cdot L$ times:

a) Choose (r, α) randomly and update
$x_{r\alpha}(t+1) = \max(0, \xi)$, where

$$\xi = \frac{J\big(1 - \sum_{\beta\neq\alpha} x_{r\beta}(t)\big) + E_{r\alpha}(t) - \mathsf{f}^{\alpha\alpha}_{rr} x_{r\alpha}(t)}{J - \mathsf{f}^{\alpha\alpha}_{rr} + T_r}$$

b) If $T_r < T^{(1)}$ and $x_{r\alpha} > \theta$, then do for all r', β: if $\mathsf{f}^{\alpha\beta}_{rr'} < 0$ and $x_{r'\beta} > \theta$ then set
$\mathsf{f}^{\alpha\beta}_{rr'}(t+1) = \mathsf{f}^{\alpha\beta}_{rr'}(t) - \Delta$ and reset
$T_r(t+1) = T_{r'}(t+1) = T^{(2)}$.

3. Set $T_r(t+1) = \eta T_r(t)$ for all r with $0 < \eta < 1$. Go to step 2 until convergence.

Step 2a) corresponds to solving the linear equation (7) $\dot{x}_{r\alpha} = 0$ independently for a randomly chosen activity $x_{r\alpha}$. In step 2b) the BDA weight modification is done if the local temperature is below the threshold $T^{(1)}$ and the two activities are greater than a threshold parameter $0 < \theta < 1$ such that an almost unambiguous labeling has already occurred for the two features. The conventional CLM algorithm is recovered by omitting step 2b). The convergence criterion was defined as

$\sum_r \left(\sum_\alpha x^2_{r\alpha}\right) / \left(L\left(\sum_\alpha x_{r\alpha}\right)^2\right) > 0.99$.

4 Application to Pentomino Tilings

The twelve pentominoes [3] are all polygons (modulo rotation and reflection) which are composed of 5 connected squares (Fig. 1). They are known to tile $6{\times}10$, $5{\times}12$, $4{\times}15$, and $3{\times}20$ rectangles. These rectangular tilings provide difficult constraint satisfaction problems which can be formulated as quadratic minimisation problems (1) and which fall into the class of NP-complete packing problems [12]. Let r denote the rth pentomino, and let $\alpha \in \{1 \ldots L_r\}$ denote the L_r possible ways of putting pentomino r onto the given rectangular field. Then we define $\mathsf{f}^{\alpha\beta}_{rr'}$ as the negative number of overlaps of pentomino r at position α with pentomino r' at position β. To allow for a rejection of single "holes", we choose $\mathsf{f}^{\alpha\beta}_{rr'} = -1$ also if a pair of pentominoes surrounds an isolated square. The task is now to find an unambiguous labeling, which forms a feasible solution and is one of the possibly many global minima with zero overlap and thus $E = 0$. To fulfil the conditions of Theorem 1 for the CLM we choose $\mathsf{f}^{\alpha\alpha}_{rr} = 1$ and take $J = 0.01 \cdot \max_{r\alpha} \sum_{r'\beta} |\mathsf{f}^{\alpha\beta}_{rr'}|$ which is sufficient to keep the deviation from the "hard" constraint (2) in the CLM small enough.

Pentomino tiling problems can nowadays be solved on standard workstations by plain backtracking algorithms within minutes. The main interest for other solution heuristics lies in the exponential scaling of the problem size for larger problems encountered e.g. in placement and packing problems in VLSI design, where mean-field-annealing has been shown to be competitive with other approaches [2]. Earlier work on solving smaller tiling problems with neural networks was based on a modified McCulloch-Pitts neuron model performing a dynamical search among discrete configura-

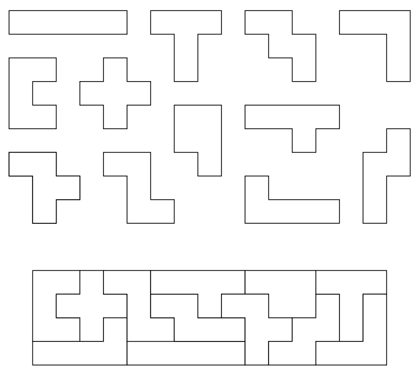

Figure 1: The twelve pentominoes and a tiling of a 4×15 rectangle.

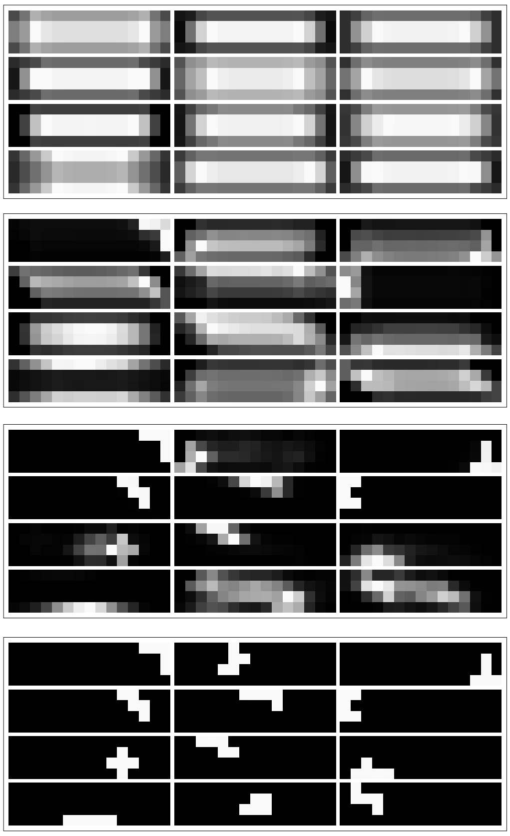

Figure 2: Graduated assignment on a 4×15 rectangle during annealing at 4 temperatures using only plain DA. Each small field shows the superposition of all the possible placements α of a single pentomino r weighted by the activity $x_{r\alpha}$ of the corresponding label. The final unambiguous output labeling has 2 overlaps.

tions of tilings [9]. Extending the algorithm by a superimposed analog representation of multiple configurations with gradual sharpening, similar to annealing, improves its performance [12].

We consider the 4×15 rectangle, for which there exist about 3×10^{24} possible alternatives of placing the twelve pentominoes if arbitrary overlaps are allowed. It is known that there exist 368 different solutions (modulo reflection and rotation). This problem is much more complex than the tiling problem considered in [9] where "only" 1.3×10^{14} combinations existed.

Figure 2 shows a time-course of annealing for both CLM or PMF. Starting from an undecided state, the assignment is gradually sharpened, until a solution (in this case suboptimal) is obtained. The intermediate stages reveal, how the interaction of all the superpositions already causes a reasonable placement of parts like the long bar-shaped and the v-shaped pentomino.

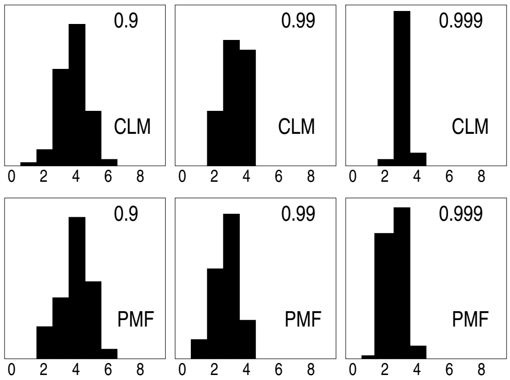

Figure 3: Performance of plain DA: Overlaps of output labelings from PMF and CLM at different annealing speeds for 1000 runs ($\eta = 0.9$ and $\eta = 0.99$) and 100 runs ($\eta = 0.999$) (bar heights normalised).

These constraints are realized by the maximum eigenvalue directions which are first explored due to the analysis in section 2. In later stages, however, certain placements which require nonlocal corrections can not be revised. Figure 4 shows histograms of the typical numbers of overlaps generated by CLM and PMF for different annealing speeds. Our PMF simulations were done by asynchronously solving the fixed point equation (4). For fast annealing CLM and PMF show similar performance with larger variation than for the slow annealing. With slow annealing the number of different output states is significantly reduced indicating more deterministic system trajectories. This effect is more pronounced for the CLM. Although slower annealing achieves on average better results for PMF, the effect saturates. In fact, even for 10^5 runs at $\eta = 0.9$ (50 sweeps per run) and 10^4 runs at $\eta = 0.99$ (500 sweeps per run) we found no solution with zero overlaps for both CLM and PMF. Note also that since annealing at $\eta = 0.99$ takes ten times more sweeps, the actual probability of obtaining a 1-overlap solution in ten $\eta = 0.9$ runs is roughly the same, indicating no advantage of slower annealing in that case.

We then performed simulations with the suggested BDA method by choosing $T^{(1)} = T_c/1000$, $T^{(2)} = T_c/10$ and modifying the weight matrix at each detected conflict by $\Delta = 1$ at a threshold of $\theta = 0.8$. Annealing was done rather fast at $\eta = 0.9$. The algorithm terminated on average after 230000 sweeps with a proper zero overlap solution (with a range from 3000 to 800000 sweeps). Although this is a considerable number of sweeps, the plain DA did not produce any zero overlap solution for altogether 5×10^6 sweeps. Throughout the iterations there were major rearrangements of single parts, however, we observed that whenever a partially valid solution is found the rearranging is temporarily restricted to the rest. To verify that the weight-changing operation is essential for the solution process we also simulated with $\Delta = 0$ where we obtained no zero overlap solution in 20 runs with 10^6 sweeps each. Another interesting observation is that, if the dynamics is restarted with an adapted weight matrix from a previous solution, the BDA terminates 10-100 times faster. This shows that somehow part of the solution has been learned by the backtracking search process which leads to an optimised weight matrix for the DA algorithm.

5 Discussion

The presented extension to deterministic annealing for constraint satisfaction problems enables the analog mean-field network to perform a search strategy in the solution space which is similar to classical backtracking algorithms. All operations within the BDA (backtracking deterministic annealing) are of local nature, involving only locally connected neurons which modify their connections due to encountered local constraint conflicts. This offers the opportunity to implement the BDA dynamics in a parallel and analog way.

It is a generic property of the formulation of constraint satisfaction problems as minimisation of quadratic energy functions that the desired global minima which satisfy all the constraints do not depend on the weight components which enforce the constraints. Therefore there is some arbitrariness involved in choosing these weights. The optimal choice depends not only on the energy function but also on the algorithm used to minimise this energy and is usually optimised by hand for a special application problem. The presented BDA network tries to automatically optimise or learn these weights during its exploration of the search space.

For the example of pentomino tilings, the BDA can obtain proper global solutions which satisfy the complex multiple constraints of the tiling problem. Although the exploration takes a considerable number of iteration sweeps and is certainly slower than plain backtracking search algorithms which were optimized for the pentomino tiling problem, our preliminary results indicate that after the optimisation a proper tiling result can be much faster obtained. We therefore consider the BDA a promising new hybrid approach which is especially relevant to problems with complex constraints where DA techniques have been succesfully applied.

Acknowledgements: H. Wersing is supported by DFG grant GK-231.

References

[1] J. R. Bitner and E. M. Reingold. Backtrack programming techniques. *Comm. ACM*, pages 651–656, 1975.

[2] T. Bultan C. Aykanat and I. Haritaoglu. A fast neural network algorithm for VLSI cell placement. *Neur. Netw.*, 11(9):1671–1684, 1998.

[3] S. W. Golomb. *Polyominoes: puzzles, patterns, problems and packings*. Princeton Univ. Press, 2nd edition, 1994.

[4] J. J. Hopfield and D. D. Tank. Neural computation of decisions in optimization problems. *Biol. Cyb.*, 52:141–152, 1985.

[5] J. J. Kosowsky and A. L. Yuille. The invisible hand algorithm: Solving the assignment problem with statistical physics. *Neur. Netw.*, 7(3):477–490, 1994.

[6] C. Peterson and B. Soderberg. A new method for mapping optimization problems onto neural networks. *Int. J. Neural Systems*, 1(1):3–22, 1989.

[7] H. Ritter. A spatial approach to feature linking. In *Int. Neur. Netw. Conf. Paris*, 1990.

[8] K. Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proc. IEEE*, 86(11):2210–2239, 1998.

[9] Y. Takefuji and K.-C. Lee. A parallel algorithm for tiling problems. *IEEE Trans. Neur. Netw.*, 1(1):143–145, March 1990.

[10] H. Wersing and H. Ritter. Feature binding and relaxation labeling with the competitive layer model. In *Proc. Eur. Symp. on Art. Neur. Netw. ESANN*, 1999.

[11] H. Wersing, J. J. Steil, and H. Ritter. A layered recurrent neural network for feature grouping. In W. Gerstner et al., editor, *Proc. ICANN*, pages 439–444, 1997.

[12] H. Yamamoto, H. Ninomiya, and H. Asai. A neuro-based optimization algorithm for rectangular puzzles. *IEICE Trans. Fundamentals.*, E 81-A(6):1113–1117, 1998.