An Algorithm for a Nonlinear Optimization Problem Using Replicator Equations

Kazuo Tsuchiya, Takehiro Nishiyama, Katsuyoshi Tsujita

Department of Aeronautics and Astronautics, Graduate School of Engineering,
Kyoto University, Kyoto, Japan.
E-mail: tsuchiya@space.kuaero.kyoto-u.ac.jp

Abstract

This paper proposes an algorithm for a nonlinear optimization problem utilizing replicator equations. The problem is to find the global minimum of a multivariate function in which each variable has a bounded feasible region. First, the feasible region of each variable is discretized and expressed as a set of nodes, and the feasible region of the problem is expressed as a set of combinations of the nodes, i.e. grid points. Then, the replicator equations are constructed with the elements which are put on the nodes. The growth rates are composed of the objective function and constraints; by increasing the parameter in the growth rate, equilibrium solutions corresponding to the grid points bifurcate successively in order, from the one having the smallest value of the function to the largest one. Lastly, an algorithm for a nonlinear optimization problem combining the deterministic annealing based on the successive bifurcation of the replicator equation and the sequential quadratic programming is developed. The algorithm is applied to the function of Fletcher & Powell, which has a lot of local minima and it is difficult to obtain the global minimum, and its performance is verified.

1 Introduction

A lot of global search methods for nonlinear optimization problem have been studied. The methods are classified into two groups, i.e. exact methods and heuristic methods. The former give the exact optimal solution with the guarantee of the optimality, but because of the computational costs, is practically impossible to be applied to large scale problems. On the other hand, the heuristic methods do not guarantee the optimality of the solution obtained, but give a good approximate solution in relatively short time. Among the heuristic methods, the optimization methods which utilize dynamical systems are studied. One of the methods is the artificial neural network model [1, 2]; as the dynamical system, the gradient vector field of the potential function composed of the objective function and the constraints is utilized. An approximate solution of the optimization problem is obtained as a stable equilibrium solution of the system. In order to improve the solution, the deterministic annealing is applied [3, 4]. We have proposed another model of the method which utilizes a dynamical system [5, 6]; as the dynamical system, a replicator equation [7] is utilized instead of a gradient vector field. The replicator equation is the equation where the derivatives of the variables are proportional to the state of the variables (the proportional coefficient is called the growth rate). The replicator equation shows successive bifurcations by increasing the value of a parameter in the growth rate. We construct the growth rate using the objective function and the constraints so that equilibrium solutions corresponding to the approximate solutions of the problem with high performance become stable earlier than those with low performance through successive bifurcations. The deterministic annealing based on the bifurcation characteristics is applied to improve the solution. In Refs. [5, 6], the proposed method has been applied to the combinatorial optimization problem and it was demonstrated that a good approximate solution can be obtained. In this paper, the proposed method is applied to a nonlinear optimization problem. The problem is to find the global minimum of a multivariate function in which each variable has a bounded feasible region. The method is applied to the function of Fletcher & Powell [8], which has a lot of local minima and therefore it is difficult to obtain the global minimum solution, and then its performance is verified.

This paper is organized as follows. In Sec. 2, the formulation of the problem is explained in detail. In Sec. 3, the dynamical system is constructed, and the stabilities of the equilibrium solutions and the bifurcation characteristics are analyzed. In Sec. 4, based on the analyses, the algorithm combining the deterministic annealing and the sequential quadratic programming [9] is proposed. In Sec. 5 results of the numerical analysis are shown. And in Sec. 6, we briefly state the conclusions and the future works.

2 Formulation of the Problem

Consider the following nonlinear continuous optimization problem

$$\min_{\{x_i\}} L(x_1, \cdots, x_N), \tag{1}$$

subject to

$$x_i^L \leq x_i \leq x_i^U \quad (i = 1, \cdots, N). \tag{2}$$

The objective function L is expressed as

$$L = \sum_m L^{(m)}(x_{i_1}, \cdots, x_{i_{n_m}}), \tag{3}$$

$$L^{(m)} = \prod_l L_l^{(m)}(x_{i_1}, \cdots, x_{i_l}), \tag{4}$$

where n_m is the number of variables included in $L^{(m)}$, and $L_l^{(m)}, L_{l'}^{(m)}$ $(l \neq l')$ do not include the same variable.

First, we reformulate the above nonlinear optimization problem as a combinatorial optimization problem. The feasible region of the variable x_i is discretized into K nodes, $x_{i1}(= x_i^L), x_{i2}, \cdots, x_{iK}(= x_i^U)$. The feasible region of the decision variables are expressed as a set of grid points $\{x_{ij}\}$. On the other hand, element S_{ij} which takes values 0 or 1 is put on the node x_{ij}. A state of the elements defines a grid point in the feasible region as

$$x_i = x_{ip(i)} \ \text{ if } \ S_{ij} = \begin{cases} 1 & (j = p(i)) \\ 0 & (j \neq p(i)) \end{cases} \quad (\forall i), \tag{5}$$

where $p = \{p(1), \cdots, p(N)\}$ is a series of integer which satisfies $1 \leq p(i) \leq K$. A new objective function $J(S_{ij})$ is defined as

$$J(S_{ij}) = \sum_m \frac{1}{n_m} \left[\prod_l \sum_{j_{i_1}=1}^{K} \cdots \sum_{j_{i_l}=1}^{K} L_l^{(m)}(x_{i_1 j_{i_1}}, \cdots, x_{i_l j_{i_l}}) S_{i_1 j_{i_1}} \cdots S_{i_l j_{i_l}} \right]. \tag{6}$$

The nonlinear optimization problem (1) is, then, reformulated as the combinatorial optimization problem,

$$\min_{\{S_{ij}\}} J(S_{ij}), \tag{7}$$

subject to

$$S_{ij} \in \{0, 1\}, \quad \sum_j S_{ij} = 1 \ (\forall i). \tag{8}$$

3 Dynamical System and Stabilities of Equilibrium Solutions

3.1 Dynamical system [5]

For the problem (7), we set the replicator equation as follows:

$$\dot{u}_{ij} = f_{ij} u_{ij}, \tag{9}$$

$$f_{ij} = (1 - u_{ij}^2) - \alpha_0 \sum_{j' \neq j} u_{ij'}^2 - \alpha_1 \lambda_{ij}, \tag{10}$$

$$(i = 1, \cdots, N; \ j = 1, \cdots, K)$$

where parameters are $\alpha_0 > 0$ and $0 \leq \alpha_1 \ll 1$, and λ_{ij} is defined as follows:

$$\frac{1}{2} \frac{\partial J(u_{i'j'}^2)}{\partial u_{ij}} \equiv \lambda_{ij} u_{ij}. \tag{11}$$

The growth rate f_{ij} is composed of three terms. The first one expresses the self activatory and inhibitory influences and leads each u_{ij}^2 to unity. The second one expresses the mutual inhibitory interactions between the elements with the same subscript i, i.e. the elements belong to the same variable x_i. The third one expresses the inhibitory influence due to the objective function and suppresses solutions with larger values of the objective function.

3.2 Equilibrium solutions and their stabilities

Equilibrium solutions of the dynamical system (9) are classified as follows:

- uniform solution:

$$u_{ij}^2 \neq 0 \quad (\forall i, j). \tag{12}$$

- feasible solution:

$$u_{ij}^2 \begin{cases} \neq 0 & (j = p(i)) \\ = 0 & (j \neq p(i)) \end{cases} \quad (\forall i). \tag{13}$$

  Each of the feasible solutions defines a grid point $(x_{1p(1)}, \cdots, x_{Np(N)})$ in the feasible region.

- transition solution:

  all the other equilibrium solutions. (14)

The results of the stability analysis of the equilibrium solutions are summarized as follows:

1. If α_0 and α_1 are sufficiently small, only the uniform solution is stable.

2. If α_1 is small and $\alpha_0 > 1$, only the feasible solutions are stable.

3. The stability condition for each feasible solution is approximately given by the following inequality:

$$\alpha_0 > 1 + \frac{\alpha_1}{N} \left(L^p - \bar{L}^p \right), \tag{15}$$

$$L^p = L(x_{1p(1)}, \cdots, x_{Np(N)}), \tag{16}$$

$$\bar{L}^p = \sum_{i=1}^{N} \sum_{m \ni i} \frac{1}{n_m} \frac{1}{K-1} \sum_{j \neq p(i)} L^{(m)}(x_{i_1 p(i_1)}, \cdots, x_{ij}, \cdots, x_{i_{n_m} p(i_{n_m})}), \tag{17}$$

where $\underset{m}{\sum_i}$ means excluding the terms not including x_i from the summation $\underset{m}{\sum}$. L^p is the value of the objective function at the grid point in the feasible region corresponding to the feasible solution, and $\bar{L}^p$ is the mean value of L^p in some neighborhood of the solution. Here we set the assumption,

1. the value of $\bar{L}^p$ is almost constant. (18)

Under the assumption (18), the second term of the right hand side of Eq. (15) is proportional to the value of L^p, and this means that the feasible solutions become stable successively in order of the value of the objective function. The assumption (18) and the stability condition (15) are checked numerically in Sec. 5.

3.3 Bifurcation characteristics

The bifurcation characteristics of the dynamical system (9) with the parameter α_0 as the control parameter are analyzed.

Consider an equilibrium solution $u_{ij} = \bar{u}_{ij}$. Solution $\bar{u}_{ij}$ changes its stability at $\alpha_0 = \bar{\alpha}_0$. Introduce vector F and matrix $D_u F$ as

$$F = [f_{11}u_{11}, \cdots, f_{1K}u_{1K}, f_{21}u_{21}, \cdots]^T \quad (19)$$

: right hand side of Eq. (9),

$$D_u F = [\partial F_{ij}/\partial u_{kl}] \quad (20)$$

: Jacobian matrix of F.

We set the following assumptions:

1. $f_{i_1 j_1}(\bar{u}, \bar{\alpha}_0) = 0$ and $\bar{u}_{i_1 j_1} = 0$ only for the pair (i_1, j_1). (21)
2. Zero eigenvalue of $D_u F$ is simple.

Under the assumption (21), there are three types of bifurcations that may occur in the dynamical system (9), i.e. saddle-node, transcritical and pitchfork bifurcations. Letting v, w be the left and right eigenvectors corresponding to the zero eigenvalue of $D_u F(\bar{u}, \bar{\alpha}_0)$, the necessary conditions for these three types of bifurcation are as follows [10]:

$$\begin{cases} \text{saddle-node:} & c_1 \neq 0, c_2 \neq 0 \\ \text{transcritical:} & c_1 = 0, c_2 \neq 0, c_3 \neq 0 \\ \text{pitchfork:} & c_1 = 0, c_2 = 0, c_3 \neq 0, c_4 \neq 0 \end{cases} \quad (22)$$

where

$$c_1 = w(D_{\alpha_0} F(\bar{u}, \bar{\alpha}_0)), \quad (23)$$
$$c_2 = w(D_u^2 F(\bar{u}, \bar{\alpha}_0)(v, v)), \quad (24)$$
$$c_3 = w(D_{\alpha_0} D_u F(\bar{u}, \bar{\alpha}_0)(v)), \quad (25)$$
$$c_4 = w(D_u^3 F(\bar{u}, \bar{\alpha}_0)(v, v, v)). \quad (26)$$

In this case, the eigenvectors are

$$v^T = w = [0, \cdots, 0, \overset{(i_1, j_1)}{\overset{\vee}{1}}, 0, \cdots, 0], \quad (27)$$

and parameters c_i are calculated as follows:

$$c_1 = \frac{\partial F_{i_1 j_1}}{\partial \alpha_0}(\bar{u}, \bar{\alpha}_0) = -\bar{u}_{i_1 j_1} \sum_{j^0 \neq j_1} \bar{u}^2_{i_1 j^0} = 0, \quad (28)$$

$$c_2 = \frac{\partial^2 F_{i_1 j_1}}{\partial u^2_{i_1 j_1}}(\bar{u}, \bar{\alpha}_0) = -6\bar{u}_{i_1 j_1} = 0, \quad (29)$$

$$c_3 = \frac{\partial^2 F_{i_1 j_1}}{\partial \alpha_0 \partial u_{i_1 j_1}}(\bar{u}, \bar{\alpha}_0) = -\sum_{j^0 \neq j_1} \bar{u}^2_{i_1 j^0} \neq 0, \quad (30)$$

$$c_4 = \frac{\partial^3 F_{i_1 j_1}}{\partial u^3_{i_1 j_1}}(\bar{u}, \bar{\alpha}_0) = -6 \neq 0. \quad (31)$$

From the results (28)–(31), only the pitchfork bifurcations may occur in the dynamical system (9). Moreover, as well as the case of Ref. [6], by increasing the value of α_0, the uniform solution finally connects with a feasible solution through pitchfork bifurcations and the feasible solution may not be the optimal solution but, in many cases, a good approximate solution.

4 Optimization Algorithm

Based on the above analyses, the optimization algorithm is proposed (Fig. 1). The solution of the deterministic annealing (Step 1–3) determines a grid point in the feasible region, and the Sequential Quadratic Programming (SQP) [9] obtains the final solution with the grid point as the initial point (Step 4).

Consequently, the final solution is the minimum of the objective function nearest to the grid point obtained by the deterministic annealing.

5 Numerical Analysis

First, numerical analysis is carried out to check the assumption (18) and the stability condition (15). The objective function used here is the function of Fletcher & Powell [8]:

$$L(\boldsymbol{x}) = \sum_{i=1}^{N} (A_i - B_i)^2, \quad (32)$$

$$A_i = \sum_{j=1}^{N} (a_{ij} \sin \alpha_j + b_{ij} \cos \alpha_j), \quad (33)$$

$$B_i = \sum_{j=1}^{N} (a_{ij} \sin x_j + b_{ij} \cos x_j), \quad (34)$$

$$-\pi \leq x_i \leq \pi, \quad (35)$$

$$a_{ij}, b_{ij} \in [-100, 100]; \ \alpha_j \in [-\pi, \pi], \quad (36)$$

Deterministic annealing

Step 1: Set α_1 and α_0 sufficiently small and compute the dynamical system (9) with random initial values to obtain the uniform solution.

Step 2: Slightly increase the value of α_0, and compute (9) with the solution of the previous iteration as the initial value to obtain the equilibrium solution.

Step 3: Go to step 2 until the feasible solution is obtained.

⇓

SQP

Step 4: Apply the SQP with the grid point obtained by the deterministic annealing as the initial point.

Figure 1: Optimization algorithm.

where a_{ij}, b_{ij}, α_j are random numbers given in Ref. [8]. This function has the obvious optimal solution $L(\boldsymbol{x}) = 0$ at $\boldsymbol{x} = \boldsymbol{\alpha}$.

With $N = 10$ and $K = 100$, the values of L^p and $\bar{L}^p$ are calculated at 100,000 grid points randomly generated. The results are shown in Fig. 2. This shows that the value of $\bar{L}^p$ is almost constant against the value of L^p.

The dynamical system (9) is calculated at many points on α_0–α_1 plane with 50 sets of random initial values. The contours of the mean value of L^p for the obtained feasible solutions are shown in Fig. 3. Figure 3 indicates that only the feasible solutions having small values of the objective function exist in the region where α_0 is small, and is consistent with the stability condition (15) qualitatively; this means that by increasing the value of α_0, the feasible solutions become stable successively in order, from the one having the smallest value of the objective function to the largest one.

Next, the performance of the proposed algorithm is verified. The function (32) is used again. The results obtained by the algorithm with $N = 2, K = 100$ are shown in Figs. 4,5. Figure 4 shows the change of the grid point on x_1–x_2 plane, where the grid point corresponding to the equilibrium solution at each iteration is given by

$$x_i = x_{ij} \quad \text{if} \quad u_{ij}^2 = \max_j u_{ij}^2 \quad (\forall i). \tag{37}$$

Figure 5 shows the change of the value of the objective function at the grid point during the process of the annealing. In Fig. 5, the value of objective function increases first, and then decreases; as

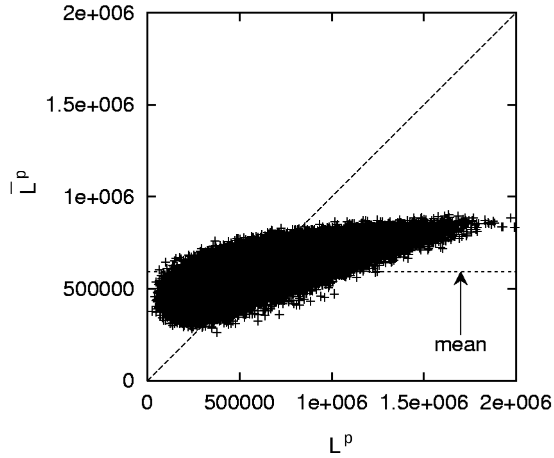

Figure 2: The value of $\bar{L}^p$ as a function of L^p.

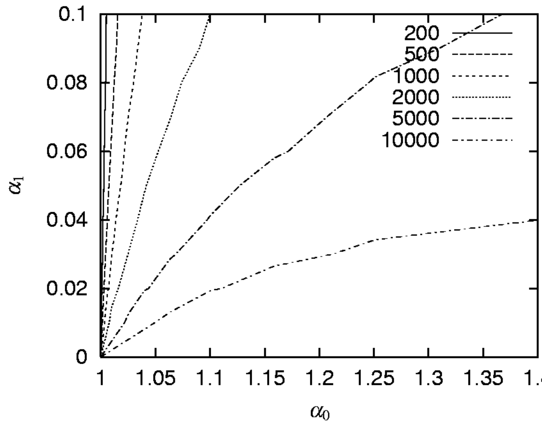

Figure 3: Contours of the mean values of the objective function L^p on α_0–α_1 plane.

shown in Fig. 4, the solution changes slightly in one of the four valleys of the objective function at the beginning of the annealing and then after jumping into another one, the solution converges to the grid point nearest to the optimal solution.

The results for $N = 30, K = 100$ are shown in Fig. 6 and Table 1. As shown in Fig. 6, the value of the objective function decreases as a whole during the annealing process and finally reaches to the optimal one by the SQP. As shown in Table 1, the value of the objective function was 673.5 for the grid point obtained by the deterministic annealing, and the computational time for the annealing was about 2 min. And the optimal solution $\boldsymbol{x} = \boldsymbol{\alpha}$ was obtained by the SQP with the grid point as the initial point. The computational time for the SQP was 2.76 sec. The distance between the grid point obtained by the annealing and the final solution was less than $2\pi/(K-1)$, the distance between the nodes, for each variable. This means that the de-

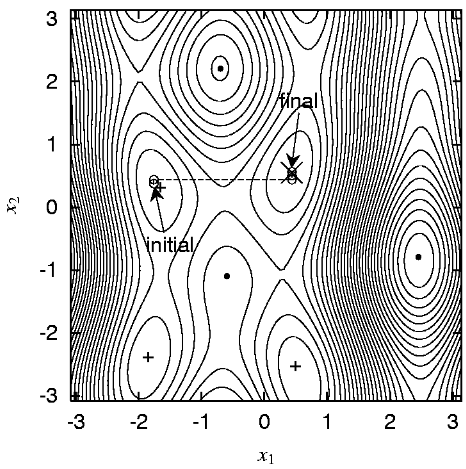

Figure 4: Change of the grid point on x_1–x_2 plane ($N = 2$, $\times$: global minimum, $+$: local minima, and $\bullet$: local maxima).

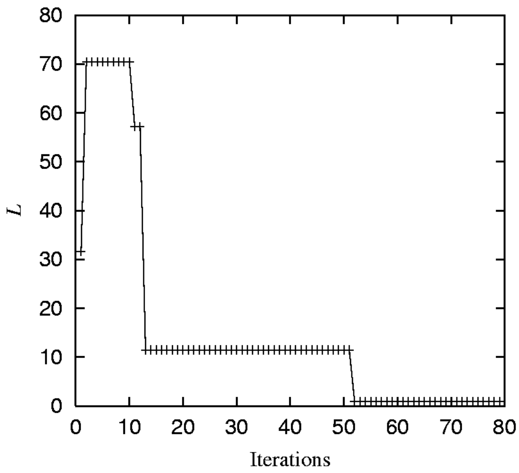

Figure 5: Change of the value of the objective function ($N = 2$ and the x-axis is the number of updates of α_0).

terministic annealing gave an almost optimal grid point, which was good enough for the initial point for the SQP.

Finally, for comparison, we applied the original SQP to the objective function (32). MATLAB Optimization Toolbox is used with 10000 sets of initial values randomly chosen from the interval $[-\pi, \pi]$. Figure 7 is the histogram of the results. The average number of function evaluations was 248, and the total computational time for 10000 trials was 44326 sec. The mean value of the objective function was $L = 965189$ and the optimal value ($L < 10^{-4}$) was obtained four times among 10000 trials. Therefore the computational time for one optimal solution is 11081.5 sec.

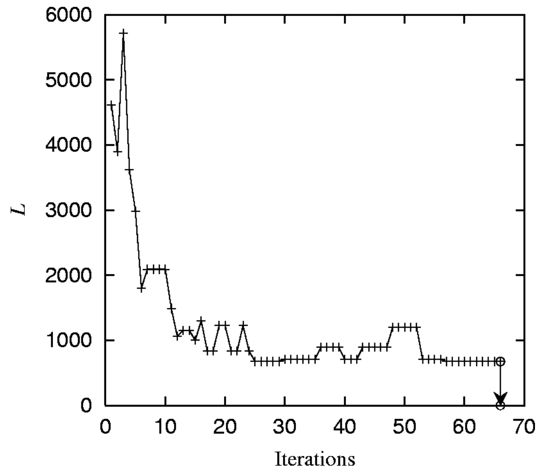

Figure 6: Change of the value of the objective function ($N = 30$, the x-axis is the number of updates of α_0 and the arrow indicates the change of the value by the SQP).

Table 1: Result of the optimization algorithm.

Deterministic annealing		
L	iterations†	CPU time*
673.5	35673	122.9 [s]
⇓		
SQP		
L	iterations‡	CPU time*
$< 10^{-6}$	352	2.76 [s]

†number of iterations of the dynamical system (9).
*on COMPAQ AlphaStation XP900.
‡number of function evaluations.

6 Conclusion

In this paper, we proposed an algorithm for a nonlinear optimization problem, in which the feasible region of the decision variables are bounded. The each feasible region are expressed as a set of finite nodes, and the feasible region of the problem is expressed as a set of combinations of the nodes, i.e. grid points. The replicator equations are constructed with the elements put on the nodes. It was shown that by increasing the value of the control parameter in the system, the equilibrium solutions corresponding to the grid points become stable successively in order, from the one with the smallest objective function value to the largest one. It was also shown that only pitchfork bifurcations occur in the dynamical system as long as the assumptions (21) hold. Based on the analyses, the algorithm combining the deterministic annealing and the Sequential Quadratic Programming (SQP) was proposed and verified numerically. The results of the numerical experiments showed that the deter-

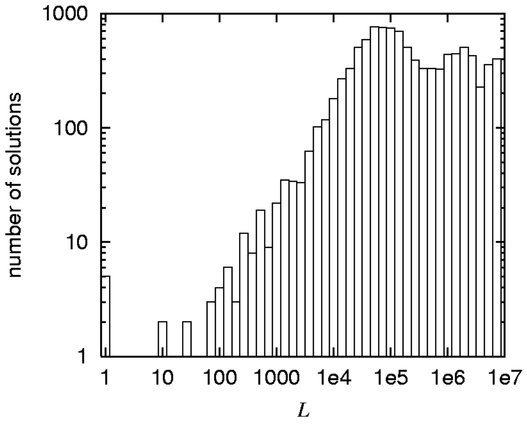

Figure 7: Results of the SQP.

ministic annealing, the first stage of the algorithm, gives the grid point sufficiently close to the optimal solution in relatively short time, and with the grid point as the initial value, the SQP gives the optimal solution.

In the formulation of the problem here, the feasible region is expressed as a set of finite grid points. To obtain the optimal grid point by enumeration, the computational complexity is about $O(K^N)$, where N is the number of the decision variables and K is the number of the nodes. But, in the proposed algorithm, the global search was carried out with computational complexity of $O(K^{N_0})$, where N_0 is the maximum number of decision variables included in $L_l^{(m)}$ of Eq. (4). If N_0 is large, the computational time increases. In this case, it is necessary to reduce the computational time. The detailed study is presently in progress.

References

[1] J. J. Hopfield and D. W. Tank. "Neural" computation of decisions in optimization problems. *Biological Cybernetics*, 52:141–152, 1985.

[2] C. R. Reeves, editor. *Modern Heuristic Techniques for Combinatorial Problems.* Blackwell Scientific Publications, 1993.

[3] G. Bilbro, R. Mann, T. K. Miller, W. E. Snyder, D. D. Van den Bout, and M. White. Optimization by mean field annealing. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems*, volume 1, pages 91–98. Morgan Kaufmann, 1989.

[4] C. Peterson and B. Söderberg. A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems*, 1:3–22, 1989.

[5] K. Tsuchiya, T. Nishiyama, and K. Tsujita. An algorithm for a combinatorial optimization problem based on bifurcation. submitted to *Neural Networks.*

[6] K. Tsuchiya, T. Nishiyama, and K. Tsujita. A deterministic annealing algorithm for a combinatorial optimization problem using replicator equations. submitted to *Physica D.*

[7] A. S. Mikhailov. *Foundations of Synergetics I*, chapter 7. Springer-Verlag, 1994.

[8] T. Bäck. *Evolutionary Algorithms in Theory and Practice*, chapter 3. Oxford University Press, New York, 1996.

[9] E. Polak. *Optimization: Algorithms and Consistent Approximations.* Springer-Verlag, 1997.

[10] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag, 1983.