Theories on the Hopfield Neural Networks

Shigeo Abe

Hitachi Research Laboratory, Hitachi, Ltd.
4026 Kuji, Hitachi, Japan

Abstract The Hopfield neural networks are well suited to solve large-scale optimization problems. But their convergence characteristics are not theoretically known.

In this paper we clarified, by an eigenvalue analysis, the conditions to converge to a vertex, a point on the edge, or an interior point of the hypercube. Then taking the traveling salesman problem as an example, we showed how to determine the weighting factors of the constraints and the objective function in the energy. Numerical calculations demonstrated that the optimal or near optimal solutions were obtained for 6, 10 and 13 cities in the traveling salesman problem.

Introduction

Until recently few efficient algorithms were known to solve a large-scale combinatorial optimization problem. But since Hopfield showed that neural networks can give the near optimal solution for the traveling salesman problem which belongs to a class of the NP complete problem, much effort has been made to apply them to several combinatorial problems. [1-14]

In the Hopfield model, the combinatorial optimization problem which minimizes a discrete objective function is converted into a continuous optimization problem which minimizes an energy function, that is a combination of constraints and an objective function. And from it a set of differential equations, each of which corresponds to the behavior of a neuron, is derived. Its integration with respect to time decreases or at least does not increase the value of the energy function. Thus we can obtain a local minimum solution where the state of each neuron is fired or not fired, namely its output is one or zero.

But because the dynamics of the neural networks have not been clarified, their application to an actual problem becomes a repetition of trial and error. For instance, there is no algorithm known to determine the values of the weighting factors for the constraints and the objective function. [8-10] There is no way to explain why a solution sometimes converges not to a vertex of a hypercube, but to a point on a surface or an interior point, *i.e.* outputs of some or all the neurons are between one and zero. Also, the solution obtained by the neural networks is not guaranteed to be optimal.

To solve these problems, first we clarify, by an eigenvalue analysis, conditions to converge to the surface of the hypercube, especially to a vertex and then show that an improper selection of initial values may lead to convergence onto an interior point. Then taking the traveling salesman problem as an example, we show how to determine the values of weighting factors. Finally, we demonstrate that an optimal or near optimal solution for the traveling salesman problem can be obtained by the proper selection of the initial values.

Problem Formulation

Let the energy function that is to be minimized be given by

$$E = 1/2\, x^t T x + b^t x \tag{1}$$

where $x = (x_1, \ldots, x_n)^t$; n-th variable vector,
$b = (b_1, \ldots, b_n)^t$; n-th input vector,

$$T = \begin{bmatrix} T_1 \\ \ldots \\ T_n \end{bmatrix} = \begin{bmatrix} T_{11} \ldots T_{1n} \\ \ldots\ldots\ldots \\ T_{n1} \ldots T_{nn} \end{bmatrix} \quad ;\ n \times n \text{ coefficient matrix,}$$

and t denotes the transpose of the matrix. In the following discussion we assume that the matrix T is symmetric, namely $T = T^t$.

To minimize E with the constraint that all x_i's be 1 or 0, is equivalent to solving the following set of differential equations[3,15] and let this be called the (1, 0) problem:

$$\begin{aligned} du/dt &= -\,\partial E/\partial x = -\,Tx - b \\ x_i &= 1/2\,(1 + \tanh\, u_i) \\ 0 &\le x_i \le 1, \text{ for } i = 1, \ldots, n \end{aligned} \tag{2}$$

The domain of vector x is the surface and the interior of the n-dimensional hypercube. The element x_i corresponds to an output of the i-th neuron, and the input to the i-th neuron is a combination of b_i and the outputs x_j multiplied by coefficients T_{ij}.

Since[15]

$$\begin{aligned} dE/dt &= \partial E/\partial x\ dx/dt \\ &= -\sum_{i=1}^{n} 2(1-x_i)\, x_i (\partial E/\partial x_i)^2 \le 0, \end{aligned}$$

namely, the energy decreases as time elapses, integration of eqs. (2) from some set of initial values leads to a local minimum solution, but no guarantee of the global minimum solution.

If all x_i's are constrained to 1 or -1 instead of 1 or 0, the set of equations corresponding to eqs. (2) is

$$\begin{aligned} du/dt &= -\,\partial E/\partial x = -\,Tx - b \\ x_i &= \tanh\, u_i \\ -1 &\le x_i \le 1, \text{ for } i = 1, \ldots, n. \end{aligned} \tag{3}$$

We call this the (1, -1) problem.

The solution of eqs. (2) or (3) may converge to any of the following:

1) a vertex of the hypercube;
2) a point on the edge of the hypercube; or
3) an interior point of the hypercube.

The solution we want is the optimal solution which converges to 1). In the following, we clarify the convergence characteristics by an eigenvalue analysis.

Eigenvalue Analysis

The solution obtained by eqs. (2) or (3) corresponds to the singular point where the derivatives of x_i's with respect to t are all zero. The stability of the singular point for the nonlinear system can be analyzed by linearizing the system and examining the eigenvalues.

First the stability of the singular points for the (1, 0) problem is considered. Eliminating u in eqs. (2) gives

$$dx/dt = -\begin{bmatrix} 2(1-x_1)x_1 & & 0 \\ & \ddots & \\ 0 & & 2(1-x_n)x_n \end{bmatrix}(Tx+b) \tag{4}$$

$$0 \le x_i \le 1, \text{ for } i = 1,..,n.$$

Thus the singular points of eq. (4) are:
(i) vertexes of the hypercube, *i.e.*, $x_i = 1$, or 0 for $i = 1,..., n$;
(ii) x that satisfies $Tx + b = 0$ and $0 \le x_i \le 1$, for $i = 1,...,n$; and
(iii) the combination of (i) and (ii).

The meaning of (iii) is as follows: Let the singular point $s = (s_1,..,s_n)^t$ satisfy

$$s_i = 1 \text{ or } 0, \text{ for } i = 1,..., k, \text{ and}$$
$$Ts + b = 0.$$

Then from eq. (4),

$$dx_i/dt = 0, \text{ for } i = 1,..., k.$$

Therefore, eq.(4) reduces to

$$\begin{bmatrix} dx_{k+1}/dt \\ \cdots\cdots \\ dx_n/dt \end{bmatrix} = -\begin{bmatrix} 2(1-x_{k+1})x_{k+1} & & 0 \\ & \ddots & \\ 0 & & 2(1-x_n)x_n \end{bmatrix}\left\{T'\begin{bmatrix} x_{k+1} \\ \cdots\cdots \\ x_n \end{bmatrix} + b'\right\} \tag{5}$$

where

$$T' = \begin{bmatrix} T_{k+1,k+1},\dots,T_{k+1,n} \\ \cdots\cdots\cdots\cdots \\ T_{n,k+1},\dots,T_{n,n} \end{bmatrix}$$

$$b' = \begin{bmatrix} b_{k+1} \\ \cdots\cdots \\ b_n \end{bmatrix} + \begin{bmatrix} T_{k+1,1},\dots,T_{k+1,k} \\ \cdots\cdots\cdots\cdots \\ T_{n,1},\dots,T_{n,k} \end{bmatrix}\begin{bmatrix} s_1 \\ \cdots \\ s_k \end{bmatrix}.$$

The dimensionality of the system is reduced from n to $n-k$. Thus all the discussion that follows holds for the reduced system.

Now the eigenvalues of the vertex of the hypercube are derived, *i.e.*, $c = (c_1,...,c_n)$, $c_i = 1$ or 0, for $i = 1,...,n$. Substitution of

$$x = y + c$$

to eq. (4) gives

$$dx/dt = dy/dt$$
$$= -2\begin{bmatrix} -y_1^2+(1-2c_1)y_1+c_1(1-c_1) & & 0 \\ & \ddots & \\ 0 & & -y_n^2+(1-2c_n)y_n+c_n(1-c_n) \end{bmatrix}\{T(y+c)+b\}. \tag{6}$$

Substitution of $c_i(1-c_i) = 0$ for $i = 1,..., n$ into eq. (6) and its linearization give

$$dy/dt = 2\begin{bmatrix} (2c_1-1)(T_1c+b_1) & & 0 \\ & \ddots & \\ 0 & & (2c_n-1)(T_nc+b_n) \end{bmatrix} y. \tag{7}$$

Thus the eigenvalues of vertex c are given by

$$\lambda_{c,i} = 2(2c_i-1)(T_ic+b_i) \text{ for } i = 1,..., n. \tag{8}$$

Therefore, vertex c is stable if

$$\lambda_{c,i} < 0 \quad \text{for } i = 1,...,n. \tag{9}$$

Let s be the singular point that satisfies (ii). In view of eq. (6), the linearized equation is given by

$$dy/dt = \begin{bmatrix} 2(s_1-1)s_1 & & 0 \\ & \ddots & \\ 0 & & 2(s_n-1)s_n \end{bmatrix} Ty. \tag{10}$$

Thus the stability of the singular point s can be determined by the eigenvalues of the coefficient matrix of eq. (10).

Now the (1, -1) problem is considered using eqs. (3). Eliminating u from eqs. (3) gives

$$dx/dt = -\begin{bmatrix} (1-x_1^2) & & 0 \\ & \ddots & \\ 0 & & (1-x_n^2) \end{bmatrix}(Tx+b) \tag{11}$$

Thus the singular points for eq. (11) are:
(i) vertexes of the hypercube, *i.e.*, $s_i = 1$, or -1 for $i = 1,..., n$;
(ii) x that satisfies $Tx + b = 0$ and $-1 \le x_i \le 1$, for $i = 1,..., n$; and
(iii) the combination of (i) and (ii).

Linearizing eq. (11) around the vertex $c' = (c_1',...,c_n')$, $c_i' = 1$, or -1, for $i = 1,..., n$, gives the following eigenvalues:

$$\lambda_{c',i} = 2c_i'(T_ic'+b_i) \text{ for } i = 1,..., n. \tag{12}$$

Therefore, the vertex c' is stable if

$$\lambda_{c',i} < 0, \text{ for } i = 1,..., n \tag{13}$$

holds.

The eigenvalues of the singular points s that satisfies (ii) are those of the following matrix:

$$\begin{bmatrix} s_1^2-1 & & 0 \\ & \ddots & \\ 0 & & s_n^2-1 \end{bmatrix} T$$

Equivalence of the (1, 0) and (1, -1) Problems

The (1, -1) problem is converted into the (1, 0) problem by substituting

$$x = 2y - \mathbf{1} \tag{14}$$

into eq. (11) where $\mathbf{1} = (1,...,1)^t$. This substitution gives

$$2dy/dt = -\begin{bmatrix} 4y_1(1-y_1) & & 0 \\ & \ddots & \\ 0 & & 4y_n(1-y_n) \end{bmatrix} [2Ty - T\mathbf{1} + b].$$

Thus

$$dy/dt = -\begin{bmatrix} 2y_1(1-y_1) & & 0 \\ & \ddots & \\ 0 & & 2y_n(1-y_n) \end{bmatrix} [2Ty - T\mathbf{1} + b]. \tag{15}$$

From eq. (8), eigenvalues of vertex c are given by

$$\lambda_{c,i} = 2(2c_i - 1)(2T_ic - T_i\mathbf{1} + b_i) \text{ for } i = 1,.., n. \tag{16}$$

Vertex c' for the (1, -1) problem corresponding to vertex c is given by

$$c' = 2c - \mathbf{1}.$$

Substituting the above equation into eq. (16) yields

$$\lambda_{c',i} = 2c'_i(T_ic' + b_i) \quad \text{for } i = 1,.., n$$

which coincides with eq. (12). Therefore, the stability of vertexes does not change even if the problem is converted from (1, 0) to (1, -1) or vice versa.

Relationships between Energy and Eigenvalues

Integrating eqs. (2) or (3) from some interior point of the hypercube gives a local minimum solution in the sense that the energy is locally minimized. But little is known about the convergence characteristics. In this section, we clarify the relationships between the energy and eigenvalues. Let vertex $c(i)$ be the same as vertex c except for the i-th element, *i.e.*,

$$c(i) = (c_1,...,c_{i-1},1-c_i,c_{i+1},...,c_n)$$

for the (1, 0) problem and

$$c(i) = (c_1,...,c_{i-1}, -c_i,c_{i+1},...,c_n)$$

for the (1, -1) problem.

In the following only the theorems for the (1, 0) problem are shown, since similar theorems can be easily derived for the (1, -1) problem.

If the energy of vertex c is lower than that of the adjacent vertex $c(i)$, the corresponding eigenvalue of vertex c is smaller than that of vertex $c(i)$. Namely:

Theorem 1 The relationship between energies and eigenvalues for adjacent vertexes c and $c(i)$ is given by

$$E_c - E_{c(i)} = (\lambda_{c,i} - \lambda_{c(i),i})/4. \tag{17}$$

Proof From eq. (1) and T being symmetric,

$$\begin{aligned} &E_c - E_{c(i)} \\ &= 1/2\ c^tTc + b^tc - 1/2c(i)^tTc(i) - b^tc(i) \\ &= 1/2c^tTc - 1/2c(i)^tTc + 1/2c^tTc(i) \\ &\quad - 1/2c(i)^tTc(i) + b^tc - b^tc(i) \\ &= 1/2(0,..,0,2c_i-1,0,..,0)\{Tc + Tc(i)\} + (2c_i-1)b_i \\ &= 1/2(2c_i-1)\{T_ic + T_ic(i) + 2b_i\}. \end{aligned}$$

Thus from eq. (8), eq. (17) holds.

The following theorem states that vertex c which has the smallest energy among the all adjacent vertexes $c(i)$, $i = 1,..,n$, becomes a stable equilibrium point when $T_{ii} = 0$ for $i = 1,..., n$.

Theorem 2 If $T_{ii} = 0$ for $i = 1,..., n$, vertex c which satisfies

$$E_c < E_{c(i)} \quad \text{for } i = 1,..., n$$

is a stable equilibrium point of eqs. (2).

Proof From eq.(8) and $T_{ii} = 0$,

$$\begin{aligned} \lambda_{c(i),i} &= -2(2c_i - 1)\ \{T_i(c_1,..,c_{i-1},1-c_i,..,c_n)^t + b_i\} \\ &= -\lambda_{c,i}. \end{aligned}$$

Also from theorem 1,

$$\lambda_{c,i} < \lambda_{c(i),i}.$$

Therefore

$$\lambda_{c,i} < 0 \text{ for } i = 1,..., n$$

holds and thus vertex c is stable.

Theorem 2 is a theoretical proof to an empirical rule [2,15] that the values of T_{ii} be set to zero so that the outputs of the neurons be one or zero. It guarantees that all the local minimum solutions are stable solutions. Thus for any local minimum solution, there is a region of initial values to converge to that solution.

The diagonal element T_{ii} and the eigenvalue $\lambda_{c,i}$ have a close relation as shown in the next theorem.

Theorem 3 If the value of T_{ii} is increased by α and that of b_i is decreased by $\alpha/2$, the eigenvalue $\lambda_{c,i}$ is increased by α, although the energy for all the vertexes remains the same. The energy for vertex c is given by

$$E_c = \Sigma'\lambda_{c,i}/4 + b^tc/2 \tag{18}$$

where Σ' means that the elements are added for $c_i = 1$. If

$$E_c > b^tc/2 \tag{19}$$

holds, vertex c is unstable.

Proof From eq. (1), the component of the energy concerning T_{ii} and b_i is

$$1/2T_{ii}x_i^2 + b_ix_i.$$

Thus if T_{ii} is increased by α and b_i is decreased by $\alpha/2$, the above component becomes

$$1/2(T_{ii}+\alpha)x_i^2 + (b_i - \alpha/2)x_i$$
$$= 1/2T_{ii}\,x_i^2 + b_ix_i + \alpha/2x_i(x_i - 1)$$
$$= 1/2T_{ii}x_i^2 + b_ix_i.$$

(The last equation is derived assuming $x_i = 1$ or 0.) Therefore, with this modification, the energies for the vertexes do not change. Since the eigenvalue $\lambda_{c,i}$ is given by

$$\lambda_{c,i} = 2(2c_i-1)(T_ic + b_i)$$
$$= 2(2c_i - 1)(T_{ii}c_i + b_i + \sum_{j\neq i} T_{ij}c_j),$$

if the values of T_{ii} and b_i are, respectively, increased by α and decreased by $\alpha/2$, the eigenvalue becomes

$$\lambda'_{c,i} = 2(2c_i - 1)(\alpha c_i - \alpha/2) + \lambda_{c,i}$$
$$= \alpha(2c_i - 1)^2 + \lambda_{c,i}$$
$$= \alpha + \lambda_{c,i}.$$

Also since

$$\{c_i/2(2c_i - 1)\}\,\lambda_{c,i} = c_i(T_ic + b_i),$$

the following equation holds:

$$\sum_{i=1}^{m}\{c_i/2(2c_i - 1)\}\lambda_{c,i} = c^t(Tc + b) = 2E_c - b^tc.$$

Thus

$$E_c = \sum_{i=1}^{n}\{c_i/4(2c_i - 1)\}\,\lambda_{c,i} + b^tc/2$$

and hence eq. (18) holds. For the stable equilibrium point,

$$\Sigma'\lambda_{c,i} < 0.$$

Thus eq. (19) is a sufficient condition for vertex c to be unstable.

If the energies are the same for the adjacent vertexes, corresponding to theorem 2, the following theorem holds:

Theorem 4 Let $T_{jj} = 0$ for $j = 1,\dots, n$. If for vertexes c and $c(i)$,

$$E_c = E_{c(i)}$$

holds and vertexes c and $c(i)$ have the smallest energy, except for each other, among the adjacent vertexes, any point on the edge connecting the vertexes c and $c(i)$ is a stable equilibrium point of eqs. (2).

Proof From the assumption,

$$\lambda_{c,i} = \lambda_{c(i),i} = 0$$
$$\lambda_{c,j} < 0$$
$$\lambda_{c(i),j} < 0 \quad \text{for } j = 1,\dots, i-1, i+1,\dots, n.$$

Let the point on the edge connecting c and $c(i)$ be expressed by

$$c' = (c_1,\dots,c_i',\dots,c_n)^t,\ 0 \le c_i' \le 1.$$

From eq. (6), the linearized equations around point c' are given by

$$dy_1/dt = \alpha_1 y_1$$
$$\dots\dots$$
$$dy_i/dt = 2c_i'(c_i' - 1)\,\Sigma\, T_{ij}y_j \qquad (c_i' \neq 1, 0)$$
$$\dots\dots$$
$$dy_n/dt = \alpha_n y_n$$

where $\alpha_j = 2(2c_j - 1)(T_jc' + b_j) < 0$ for $j = 1,\dots, i-1, i+1,\dots,n$.

Thus

$$y_j = \beta_j e^{\alpha_j t}$$

and

$$y_i = \Sigma k_j e^{\alpha_j t}.$$

Therefore, the point c' is stable when $0 < c_i' < 1$. We can similarly prove stability for $c_i' = 1$ and 0.

Theorems 2 and 4 tell us that for $T_{ii} = 0$, $i = 1,\dots, n$ the solution, when converged to the surface of the hypercube, converges either to

(1) a vertex, or

(2) a point on the edge connecting the vertexes or the space less than n dimensions.

If the solution converges to a surface of the hypercube, any vertex within the space can be selected as a solution, since the energies of any vertexes are the same. When $T_{ii} \neq 0$, a point on the surface of the hypercube can be a stable solution. The following theorems explain this.

Theorem 5 Assume that

$$E_c < E_{c(i)},\ T_{ii} > 0 \text{ and } \lambda_{c(i),i} > \lambda_{c,i} > 0.$$

There is a singular point on the edge connecting c and $c(i)$ between c_i and 1/2, and the singular point is stable when eqs. (2) reduce to the edge.

Proof From eq. (4), the equation which defines the i-th eigenvalue is

$$dy_i/dt = -2y_i(1-y_i)\{T_i(c_1,..,y_i,..,c_n)^t + b_i\}. \qquad (20)$$

Thus the singular points for this system are $y_i = 0$, 1, and

$$T_{ii}y_i = -(\sum_{j\neq i} T_{ij}c_j + b_i). \qquad (21)$$

From the relation of the eigenvalues,

$$-2(2c_i - 1)\{T_{ii}(1 - c_i) - T_{ii}y_i\} > 2(2c_i-1)\{T_{ii}c_i - T_{ii}y_i\} > 0.$$

Thus from $T_{ii} > 0$

$$(2c_i - 1)(c_i - y_i) > 0$$

and

$$(2c_i - 1)(1 - 2y_i) < 0.$$

Therefore

$$1/2 < y_i < 1 \quad \text{for } c_i = 1$$

and $\quad 0 < y_i < 1/2 \quad \text{for } c_i = 0.$

Since the coefficient of y_i in eq. (20) is negative, the singular point satisfying eq. (21) is stable within the framework specified by eq. (20).

Theorem 6 Let

$$E_c = E_{c(i)}, \text{ and } T_{ii} > 0.$$

Then $(c_1,..,c_{i-1},1/2,c_{i+1},...,c_n)$ is a singular point, and it becomes stable when eqs. (2) reduce to the edge connecting c and $c(i)$.

Proof From eq. (8),

$$\lambda_{c,i} = \lambda_{c(i),i} = 2(2c_i-1)(T_ic + b_i) = -2(2c_i - 1)(T_ic(i) + b_i).$$

Thus

$$1/2T_{ii} = -\Sigma_{i\neq j}(T_{ij}c_j + b_i) \tag{22}$$

and $\quad \lambda_{c,i} = \lambda_{c(i),i} = 2(2c_i - 1)(T_{ii}c_i - 1/2T_{ii}) = T_{ii}.$

Substituting eq. (22) into eq. (21) gives $y_i = 1/2$. Similar to the proof of theorem 5, the singular point is shown to be stable when eqs. (2) reduce to the edge connecting vertexes c and $c(i)$.

Theorems 5, and 6 show the possibility of the solution to converge to a point on the edge. If that happens, the theorems tell that the vertex which has a lower energy is the one which is nearer to the solution.

Theorem 7 Let eqs. (2) have a stable vertex c and the singular point s given by $Ts + b = 0$ satisfy $0 \le s_i \le 1$ for $i = 1,...,n$. Then $E_c \le E_s$ holds where the equality holds for $s = c$.

Proof From eq. (18),

$$E_c = \Sigma'\lambda_{c,i}/4 + b^tc/2.$$

Also from $Ts + b = 0$,

$$E_s = 1/2s^tTs + b^ts = 1/2b^ts.$$

Therefore

$$E_c - E_s = 1/4\Sigma'\lambda_{c,i} + 1/2b^t(c-s).$$

As

$$b^t(c-s) = -s^tT(c-s) = -s^t(Tc + b) = -\Sigma\{s_i/2(2c_i-1)\}\ \lambda_{c,i},$$

the following equation holds:

$$E_c - E_s = 1/4\ \Sigma'\{1-s_i/(2c_i-1)\}\lambda_{c,i} - 1/4\Sigma''\{s_i/(2c_i-1)\}\ \lambda_{c,i}$$

where Σ' is added for $c_i = 1$ and Σ'' is added for $c_i = 0$. As $\lambda_{c,i} < 0$ and $0 \le s_i \le 1$,

$$E_c - E_s = 1/4\Sigma'(1 - s_i)\lambda_{c,i} + 1/4\Sigma''s_i\lambda_{c,i} \le 0.$$

Since there is no singular point within the hypercube except the singular point s, theorem 7 suggests that it becomes an unstable equilibrium point when there are stable vertexes, and there is a region around s where the solution converges to any stable vertex.

Convergence to an Interior Point

Even if T_{ii} are set to zero, a mal-selection of initial values may lead eqs. (2) to converge to an interior point of the hypercube. The following two theorems explain this.

Theorem 8 Let the absolute values of all the elements of $b + 1/2T\mathbf{1}$ be the same, and the singular point s given by $Ts+b = 0$ satisfy

$$2s_i - 1 = tanh\ \alpha\ (b + 1/2T\mathbf{1})_i \text{ for } i = 1,..., n$$
$$\text{and } \alpha < 0, \tag{23}$$

eqs. (2) converge, by the ***Euler*** method, from the initial values $1/2\ \mathbf{1}$ to s with the absolute values of x monotonically increasing for $-\alpha > \Delta t$.

Proof Since $x(0) = 1/2\ \mathbf{1}$, $u(0) = 0$ holds. By the ***Euler*** method,

$$u(i+1) - u(i) = -\Delta t(Tx(i) + b)$$
$$x(i) = 1/2\ (\mathbf{1} + tanh\ u(i)) \tag{24}$$

where $tanh\ u(i) = (tanh\ u(i)_1,..., tanh\ u(i)_n)^t$.

Thus $u(1) = -\Delta t(1/2T\mathbf{1} + b)$.

From the assumption, absolute values of the elements of $u(1)$ are the same. Thus the absolute values of the elements of $2x(1) - \mathbf{1} = tanh\ u(1)$ are the same and the signs of $u(1)_i$ and $(2x(1) - \mathbf{1})_i$ are the same for $i = 1,..., n$. Similarly, we can prove that the absolute values of the elements of $u(i)$ are all the same.

Multiplying T from the left hand side of eq. (23) gives

$$Ts = 1/2T\mathbf{1} + 1/2\ T\ tanh\ \alpha(b + 1/2T\mathbf{1}).$$

Thus

$$Ts + b = 1/2T\mathbf{1} + b + 1/2T\ tanh\ \alpha(b + 1/2T\mathbf{1}) = 0$$

or equivalently,

$$u(1) - (\Delta t/2)T\ tanh\ \{(-\alpha/\Delta t)u(1)\} = 0. \tag{25}$$

This means that the absolute values of elements of ***Ttanh*** $u(1)$ are the same, and signs of the elements are the same as those of the corresponding elements of $u(1)$.

From eq. (24),

$$u(i + 1) = u(i) - \Delta t(Tx(i) + b)$$
$$= u(i) - \Delta t(1/2T\ tanh\ u(i) + 1/2T\mathbf{1} + b)$$
$$= u(i) + u(1) - (\Delta t/2)\ T\ tanh\ u(i). \tag{26}$$

Now let

$$f(x) = x - (\Delta t/2)\ T\ tanh\ x,$$
$$x = (x_1,..., x_n)^t$$
$$|x_i| = |x_j|, \text{ for } i, j = 1,..., n. \tag{27}$$

From the last constraint in eq. (27), $f(x)_i$ is equivalent to

$$f(x)_i = x_i - a_i \tanh x_i$$

where a_i is constant.

Therefore, eq. (26) becomes

$$u(i+1) = u(1) + f(u(i)). \tag{28}$$

From eq. (25) and $-\alpha > \Delta t$, for $|x| \geq |u(1)|$, the signs of the elements of $f(x)$ are the same as those of the corresponding elements of x. Therefore,

$$|u(2)| = |u(1) + f(u(1))| \geq |u(1)|.$$

Since $|f(x_2)| \geq |f(x_1)|$ holds for $|x_2| > |x_1| > |u(1)|$,

$$|u(3)| = |u(1) + f(u(2))| \geq |u(2)|.$$

Likewise

$$|u(i+1)| \geq |u(i)| \geq ,..., \geq |u(1)|.$$

Thus the absolute values of the elements of $u(i)$ increase monotonically as i increases. Monotonic increase guarantees that the signs of the elements of

$$u(i+1) - u(i) = u(1) - \Delta t/2 \tanh u(i)$$

are the same as those of the corresponding elements of $u(1)$. Namely,

$$|u(1)| \geq |\Delta t/2\ T \tanh u(i)|.$$

Thus from the monotonic increase of the absolute values of the elements of $u(i)$ and the upper bound,

$$|u(1)| = |\Delta t/2\ T \tanh u(i)|$$

holds as $i \to \infty$ which means that $x(i)$ converges to the solution of $Tx + b = 0$.

The assumption for theorem 8 is very special. There is a more general case where the solution converges to an interior point, as follows:

Theorem 9 Let a set of subscripts be defined as follows:

$$N_1 \cup ... \cup N_k = \{1, 2,..., n\}$$
$$N_1 \cap ... \cap N_k = \emptyset,$$

and for any $i, j \ni N_l$

$$w(0)_i = w(0)_j,$$
$$(1/2T\mathbf{1} + b)_i = (1/2T\mathbf{1} + b)_j, \text{ and}$$
$$(Tw(0))_i = (Tw(0))_j$$

or

$$w(0)_i = - w(0)_j$$
$$(1/2T\mathbf{1} + b)_i = - (1/2T\mathbf{1} + b)_j, \text{ and}$$
$$(Tw(0))_i = - (Tw(0))_j. \tag{29}$$

Then if the singular point s given by $Ts + b = 0$ satisfies $0 \leq s_i \leq 1$ for $i = 1,..., n$ and the initial vector $x(0)$ is selected to be $w(0) + 1/2\ \mathbf{1}$, the n-dimensional system given by eqs. (2) reduces to the k-dimensional system. There is also a case where the singular point s for the reduced system becomes stable even if the original singular point is unstable.

Proof Substituting

$$x = w + 1/2\ \mathbf{1}$$

into eq. (4) gives

$$dw/dt = -\begin{bmatrix} 2(1/2 - w_1)(1/2 + w_1) & & 0 \\ & \ddots & \\ 0 & & 2(1/2 - w_n)(1/2 + w_n) \end{bmatrix} (Tw + 1/2T\mathbf{1} + b). \tag{30}$$

If $w = w(0)$, according to the assumption, absolute values of all the w_i, for $i \in N_l$, are the same for $t \geq 0$. Thus the behavior of w_i, $i \in N_l$ is represented by the behavior of one w_i. Selecting $i \in N_l$ as the representative of N_l, the coefficient T'_{ii} of w_i is given by

$$T_{ii}' = T_{ii} + \sum_{\substack{j \in N_l \\ w_i = w_j}} T_{ij} - \sum_{\substack{j \in N_l \\ w_i = -w_j}} T_{ij}.$$

Thus even if $T_{ii} = 0$ for the original system, there is a case where T'_{ii} becomes positive. Therefore, interior point may become stable and all the vertexes become unstable.

As we show in the following, theorem 9 holds for the traveling salesman problem. To avoid degeneration, the initial values must be selected as

$$x(0) = 1/2\ \mathbf{1} + w(0)$$

where $|w(0)|_i \neq |w(0)|_j$, $i \neq j$, $i, j \in N_l$, $l = 1,..., k$. (31)

Convergence to the Optimum Solution

So far we have clarified the convergence characteristics; convergence to the optimum solution still remains. But just now what we can prove is as follows:

Theorem 10 If the singular points given by (ii) and (iii) (applying to eq. (4)) do not exist in the hypercube, there are no more than two isolated local minimum solutions. Namely, the global minimum solution is obtained setting arbitrary initial values within the hypercube.

Proof Assume there are two isolated local minimums. Since there are no singular points within the hypercube, any point within the hypercube converges to either of the two singular points. This means that the hypercube is separated into two regions. Since there are no singular points within the hypercube, the intersection of the separation surface and the hypercube becomes a singular point, which contradicts the assumption.

The above theorem guarantees the global convergence to the optimum solution, but actually it seems very rare that the assumption of the theorem holds.

Uesaka[15] conjectured that if the initial values are selected nearer to the center of the hypercube, the more likely that the optimum solution can be obtained. This seems to be a good conjecture at present.

Traveling Salesman Problem

Problem formulation

The traveling salesman problem is to travel to n cities within the minimum distance with the constraint that all the cities must be visited only once. Hopfield assigned n neurons for each city and stated that if the output of i-th neuron of the city is one, the city is visited at the i-th order. He defined the energy function as follows:

$$\begin{aligned} E = {} & A/2 \sum_{x} \sum_{i} \sum_{j \neq i} V_{xi} V_{xj} + B/2 \sum_{i} \sum_{x} \sum_{x \neq y} V_{xi} V_{yi} \\ & + C/2 \Big(\sum_{x} \sum_{i} V_{xi} - n \Big)^2 \\ & + D/2 \sum_{x} \sum_{y \neq x} \sum_{i} d_{xy} V_{xi} (V_{y,i+1} + V_{y,j-1}) \end{aligned} \tag{32}$$

where V_{xi} : i-th neuron for city x, and $0 \le V_{xi} \le 1$,
d_{xy}: distance between city x and city y, and
A, B, C: weighting factors.

The first term of the energy function specifies that for city x, the number of subscripts i that satisfy $V_{xi} = 1$ is at most one. The second term specifies that for subscript i, the number of cities x that satisfy $V_{xi} = 1$ is at most one. And the third term specifies that the total number of neurons that fire is n. The first three terms are constraints. The constraint specified by the third term is relatively weak, since the third term allows that for city x, $V_{xi} = 1$ for $i = 1, \ldots, n$ and the remaining $V_{yi} = 0$, which contradicts the first two terms. [10] Therefore instead of eq. (32) we introduce the following energy function:

$$\begin{aligned} E = {} & A/2 \sum_{x} \Big(\sum_{i} V_{xi} - 1 \Big)^2 + B/2 \sum_{i} \Big(\sum_{y} V_{yi} - 1 \Big)^2 \\ & + D/2 \sum_{x} \sum_{y \neq x} \sum_{i} d_{xy} V_{xi} (V_{y,i+1} + V_{y,i-1}). \end{aligned} \tag{33}$$

The first and the second terms carry the constraint that $V_{xi} = 1$ holds for one x or i while varying them from 1 to n. The weighting factors A, B, and D can be determined as follows: Weighting factors A and B are for constraints. And these constraints must work equally. If for a solution that satisfies the constraints one output of the neurons changes from one to zero, the energies corresponding to the first and the second constraints increase by $A/2$ and $B/2$. Thus for the two constraints to work equally,

$$A = B \tag{34}$$

must hold.

In order for the solution that satisfies the constraints to be a local minimum one, the weighting factor D must be selected so that their energy must be lower than that of the solutions which do not satisfy constraints. If for the solution that satisfies the constraints, one output of the neurons changes from one to zero, the energy corresponding to the third term decreases at a maximum by

$$2D \max_{x,\, y} d_{xy}$$

whereas the energy corresponding to the first and the second terms increases by A. Thus to obtain a solution which satisfies the constraints,

$$A > 2D \max_{x,\, y} d_{xy} \tag{35}$$

must hold.

Changing $V_{xi}{}^2$ into V_{xi}, we can obtain T and b defined in eq. (1) as follows:

$$T = \left[\begin{array}{ccccc|ccccc|c} 0 & A & \cdots & \cdots & A & A & D_{xy} & 0 & \cdots & 0 \; D_{xy} & \\ A & 0 & \cdots & \cdots & A & D_{xy} & A & D_{xy} & 0 \cdots & 0 & \\ & & \ddots & & & & & \ddots & & & \\ A & \cdots & \cdots & A & 0 & D_{xy} & 0 & \cdots & D_{xy} & A & \\ \hline A & D_{xy} & 0 & \cdots & D_{xy} & 0 & A & \cdots & \cdots & A & \\ & \ddots & & & & & \ddots & & & & \\ D_{xy} & 0 & \cdots & D_{xy} & A & A & \cdots & \cdots & A & 0 & \\ \hline & & & & & & & & & & \end{array} \right]$$

$$b = -(A, \ldots\ldots, A)^t \tag{36}$$

where $D_{xy} = D d_{xy}$.

Seen from T and b, we can easily find that if we select initial values as

$$V_{xi} = 1/2\, \mathbf{1} + w_x \quad \text{for } i = 1, \ldots, n \tag{37}$$

theorem 9 holds, and V_{xi} for $i = 1, \ldots, n$ behaves in the same way, and the system reduces from n^2 dimensions to n dimensions. Since the diagonal elements of the reduced matrix of T' become $(n-1)A$, the interior singular point may become stable. Therefore initial values must be selected so that at least values for V_{xi} for $i = 1, \ldots, n$ are different.

From eqs. (8) and (36), the eigenvalues of vertex c that satisfy the constraints are

$$\begin{aligned} \lambda_{c,i} & = 2(D_{xy} + D_{xz}) - 2A && \text{for } c_i = 1, \text{ and} \\ & = -2(D_{xy} + D_{xz} + A), && \text{or} \\ & = -2(D_{xy} + A) && \text{for } c_i = 0. \end{aligned} \tag{38}$$

Thus eq. (35) is necessary for the solution which satisfies the constraints to be stable.

Numerical calculations

Based on eq. (33), we solved the traveling salesman problem for 6, 10 and 13 cities as shown in Table 1. The initial values were selected according to eq. (31). The absolute values of the elements of $w(0)$ were selected as small as possible. That meant if the values were so small that the solution converged to an interior point, they were increased by a small increment until the solution converged to a vertex point. For 6, and 13 cities, the optimal problem solutions were obtained. For the 10-city traveling salesman problem, the solution was not optimal, but was quite near the optimal solution.

There was almost no problem in doing the computer simulations. Even if some problem arises, the theories suggest proper actions to take. But as for the convergence to the optimal solution, further investigation is still needed.

Table 1 Solutions for the traveling salesman problem

No. of cities	No. of solutions	Obtained distance solution	Optimum distance solution
6	60	104	104
10	181,440	159	152
13	239,500,800	835	835

Conclusions

The convergence characteristics of the Hopfield neural networks were clarified and the conditions to converge to a vertex, a point on the edge or an interior point of the hypercube were clearly stated.

Then taking the traveling salesman problem as an example, determination of the weighting factors of the constraints and the objective function in the energy were shown. Numerical calculations demonstrated that optimal or near optimal solutions were obtained for 6, 10 and 13 cities in the traveling salesman problem.

Acknowledgements

We wish to thank Dr. Hirasawa and Dr. Kawakami of Hitachi Research Laboratory and the members of the HRL Neural Project for their helpful discussions and encouragement.

References

1) J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proc. Natl. Acad. Sci. USA, Vol. 79, pp. 2554-2558 April 1982.
2) J.J. Hopfield, "Neurons with Graded Response Have Collective Computational Properties like those of Two-State Neurons," *ibid.*, Vol. 81, pp. 3088-3092, May 1984.
3) J.J. Hopfield and D.W. Tank, "Neural Computation of Decisions in Optimization Problems, " Biological Cybernetics, Springer-Verlag Vol. 52, pp. 141-152, 1985.
4) R.P. Lippmann, "An Introduction to Computing with Neural Nets," IEEE ASSP Magazine, pp.4-22, April 1987.
5) B. W. Lee, B. J. Sheu, "An Investigation on Local Minima of Hopfield Network for Optimization Circuits," Proc. ICNN-88, pp45-51, July 1988.
6) D. J. Burr, "An Improved Elastic Net Method for the Traveling Salesman Problem," *ibid.*, pp69-76.
7) G. J. Hueter, "Solution of the Traveling Salesman Problem with an Adaptive Ring," *ibid.*, pp85-92.
8) H. Szu, "Fast TSP Algorithm Based on Binary Neuron Output and Analog Neuron Input Using the Zero-Diagonal Interconnect Matrix and Necessary and Sufficient Constraints of the Permutation Matrix, *ibid.*, pp259-266.
9) S. U. Hegde, J. L. Sweet, W. B. Levy, "Determination of Parameters in a Hopfield/Tank Computational Network," *ibid.*, pp 291-298.
10) D. E. Van den Bout, T. K. Miller, "A Traveling Salesman Objective Function that Works," *ibid.*, pp 299-303.
11) E. Wacholder, J. Han, R. C. Mann, "An Extension of the Hopfield-Tank Model for Solution of the Multiple Traveling Salesman Problem," *ibid.*, pp305-324.
12) J. Ramanujam, P. Sadayappan, "Optimization by Neural Networks," *ibid.*, pp325-332.
13) R. D. Brandt, Y. Wang, A. J. Laub, S. K. Mitra, "Alternative Networks for Solving the Traveling Salesman Problem and the List-Matching Problem," *ibid.*, pp333-340.
14) Y. S. Foo, Y. Takefuji, "Integer Linear Programming Neural Networks for Job-Shop Scheduling," *ibid.*, pp341-348.
15) Y. Uesaka, "On the Stability of Neural Network with the Energy Function Induced from a Real-Valued Function of Binary Variables, " IEICE Technical Report, PRU-88-6(in Japanese), June 1988