

(43) ニューロ・オプティマイザの開発に向けて

NEURO-OPTIMIZER, BASIC CONCEPT AND STUDIES

岸 光 男*

Mitsuo KISHI,

鈴木 泰三**

Taizo SUZUKI

Most of practical design optimizations are combinatorial optimization problems on account of the industrial standard or of the inherent characteristic of the design variable. Some combinatorial optimization problems can be programmed and solved on artificial neural network system minimizing the Lagrangean, however, which must be quadratic. We propose a design support system named NEURO-OPTIMIZER. The basic concept is as follows: i) The system attains good (not necessarily best) solutions by using advanced neuron model. ii) The simulated annealing method is introduced to escape from local minima. In this paper we investigate the number representation for the design variable. Numerical examples for a nonlinear combinatorial optimization problem are provided to illustrate basic properties of the NEURO OPTIMIZER.

Key words : Neural network, Combinatorial optimization, Nonlinear programming, Number representation, Simulated annealing

1. まえがき

CAE(Computer Aided Engineering)の発展により、数値解析に立脚した設計(Design by Analysis)が指向されていく。解析に止まらず設計プロセス中でコンピュータが支援しうるステップは多い。設計手順が類型化されている場合には、最も創造的な概念設計のステップでさえエキスパート・システム化することができる。設計とは仕様を満たす解(設計諸元や属性)の探索・決定である。最終的な意思決定はデザイナーに帰するが、詳細設計のステップで設計代替案中に具体的な属性値を設定する作業にはコンピュータからの支援が可能である。すなわち、何らかの最適化機能を用いての設計諸元の決定である。

実際の機械・構造システムの設計問題では、設計諸元(設計変数)は工業規格やその変数固有の性質により連続量ではなく離散量として取扱う必要がある場合が多い。したがって、その設計問題の最適化は組合せ最適化問題となる。組合せ最適化問題に対しては分岐限定法を中心とする種々の手法が存在している。それらの手法では、対象とする問題固有のヒューリスティックス(発見的)な知識を利用しない限り処理時間は問題規模の指数オーダーで増加していくことから、大規模システムに対して簡便・実用的であるとはいえない。

最近、人工的な神経回路網(ニューラル・ネット)による並列情報処理に適した組合せ最適化の近似解法が提案され^{1), 2)}, アルゴリズムの単純さおよび処理時間の高速性が見込まれるという点で注目されている。ニュー

* 工博 大阪府立大学講師 工学部船舶工学科

** 大阪府立大学大学院生 工学研究科船舶工学専攻

ニューラル・ネットによる近似解法では、最適化問題中の制約条件にペナルティ乗数（Lagrange乗数）をかけ、それを目的関数と和してLagrange関数を生成する。そして、ニューラル・ネットのエネルギー最小化の原理によりそのLagrange関数を最小化することによって準最適解を見つけ出す。ただし従来の研究では、i) Lagrange乗数の値を経験的に与えていること、ii) 目的関数および制約条件が設計変数（制御変数）に関して2次までの問題に限定される、という点で限界があった。

i)の点に対して、著者らは数理計画法の制約式変換法に準じてLagrange乗数を更新することを試み、その有効性を確認している³⁾。たとえば、Fig. 1の船体中央横断面に対して設計規則からの縦強度制約のもと構造断面積の最小化を試みた。実際の設計では規格鋼材が使用されることから、各縦強度部材は局部強度を満足するそれぞれ10個の候補部材の中から選定するものとした。グローバルな最適解を探索する目的から、ニューラル・ネットに確率的な外乱を加え、やきなまし法⁴⁾でその外乱を調整した。Fig. 2は、3通りの温度パラメタに対するニューラル・ネットのエネルギー変化の様子を示している。温度パラメタが大きい場合、任意に抽出した2万個の許容解（Fig. 3参照）のいずれよりも優れた解をニューラル・ネットは見つけ出している。

ii)の点は、ニューラル・ネットの最適設計への応用を考える上で大きな障害となる。先に示した船体中央横断面の最適設計問題でも、そのために候補部材の違いによる部材重心位置の変化を考慮することができていない。ニューラル・ネットによる組合せ最適化問題の近似解法をより一般的な非線形計画問題に適用可能とさせる試み⁵⁾がなされているが、まだ実用性を見込める成果は出されていない。著者らは、シナプス荷重の概念を棄却し、最急降下法によるLagrange関数の最小化にニューロンの入出力モデルを組込むことを考える。すなわち、組合せ最適化問題に対する汎用的な準最適化ツールとしてのニューロ・オブティマイザの提案である。本研究では特に、ニューロンによる離散変数の数値表現法がニューロ・オブティマイザの挙動に及ぼす影響について調べる。

2. ニューロ・オブティマイザの基本原理解

2.1 エネルギー最小化の原理

Hopfieldモデルに代表されるニューラル・ネットのエネルギー最小化の原理は、ニューロンの入出力モデルが介在すること以外は最急降下法と全く同形である。したがって、シナプス荷重の概念を変容させれば、その原理は一般的な非線形エネルギー関数 E に対しても拡張される。まず、時刻 t におけるニューロン i の状態 $V_i(t)$ 、入力 $U_i(t)$ に対して次の状態方程式を導入する。

$$V_i(t) = \phi(U_i(t)) \quad \dots\dots\dots(1)$$

$$dU_i/dt = -\partial E/\partial V_i \quad \dots\dots\dots(2)$$

入出力関数 ϕ は 0 から 1 までの値をとる単調増加関数であり、たとえば、

$$\phi(u) = [1 + \tanh(u/u_0)]/2 \quad \dots\dots\dots(3)$$

(u_0 は定数) のようなシグモイド型の関数を用いられる。エネルギー E が各ニューロンの状態 V の関数であるならば、(1)～(3)式を用いると次式が容易に導出される。

$$dE/dt = \sum_i (\partial E/\partial V_i) (dV_i/dt) = \sum_i (\partial E/\partial V_i) (dV_i/dU_i) (dU_i/dt) = -\sum_i \phi'(\partial E/\partial V_i)^2 \leq 0 \quad \dots\dots(4)$$

ここにニューロンの状態が最終的に 0 または 1 (0-1ニューロン) となるような制約を組込めば、それは Hopfieldモデルと同じく組合せ最適化問題の近似解法としての性質をもつことになる。

2.2 確率論的アプローチ

2.1のエネルギー最小化の原理は、各ニューロンの状態がエネルギーの大域的な最小点に遷移することを保証

するものではない。局所的な最小点からの脱出方法として、やきなまし法をベースとする種々の確率論的アプローチが考え出されている。その一つとして、(2)式に確率的なノイズを付加する方法がある。すなわち、

$$dU_i/dt = -\partial E/\partial V_i + e \quad \dots\dots\dots(5)$$

ここで、 e は平均 0、分散 $\sigma^2(=\alpha T^2$ ： α は正の定数、 T はやきなまし温度)の正規確率変数である。やきなまし温度 T は、たとえば $T(t)=T_0/\ln(1+t)$ (T_0 は温度パラメタ)なるスケジュールで冷やしていく。以上のアプローチは非常に簡便ではあるが、必ずしもニューロンの最終状態を 0 または 1 にすることができないという欠点がある。別な確率論的アプローチとしてボルツマン・マシン⁶⁾がある。上述のアプローチに比較すると、ボルツマン・マシンでの計算手順は複雑であるが、各ニューロンの状態は常に 0 または 1 となることが保証される。計算時間を短くするために、確率論的アプローチを決定論的アプローチに変換する平均場近似アニーリング法⁷⁾と呼ばれる手法も提唱されている。そこでは、ニューロンの状態の平均値が統計熱力学の手法を用いて解析的に計算される。

3. 数値表現法

Hopfieldモデルでは、0-1ニューロンによる離散変数の数値表現法がニューラル・ネットの挙動を支配する重要な要素であった^{2), 3)}。一般的な非線形問題を取扱うニューロ・オプティマイザでは、数値表現法に対してより大きな注意を払う必要がある。

3.1 分散的表現

ニューラル・ネットワークの最大の特徴は分散的情報処理にある。ニューロンによる数値表現でも、冗長性のある分散的な表現が望ましい。たとえば、次のような規則的離散変数の集合 $S_r=\{\gamma, \gamma+\beta, \dots, \gamma+i\beta, \dots\}$ に対しては、

$$X = \beta \sum_i V_i + \gamma \quad \dots\dots\dots(6)$$

なる表現が用いられる。この場合、表現上は各ニューロンはお互いに独立である。

3.2 局所的表現

不規則的離散変数の集合 $S_{ir}=\{a_1, a_2, \dots, a_i, \dots\}$ に対しては、

$$X = \sum_i a_i V_i, \quad \sum_i V_i = 1 \quad \dots\dots\dots(7)$$

なる表現が考えられる。この場合、各ニューロンはお互いに排他的である。

3.3 中間的表現

船体中央横断面の縦強度設計で示されているように、Hopfieldモデルでは局所的表現でも準最適解が得られる。しかしながら、局所的表現は冗長性がなくニューラル・ネットの特長を生かすことにならない。そこで、不規則的離散変数の集合 S_{ir} に対して、分散的表現と局所的表現との中間的な表現が望まれる。たとえば、

$$X = a_1 + \sum_i (a_{i+1}-a_i)V_i, \quad V_i \geq V_{i+1} \text{ for all } i \quad \dots\dots\dots(8)$$

なる表現が考えられる。ただし、 $a_{i+1} > a_i$ (for all i)である。この場合、各ニューロンはお互いに従属的である。

その他に、不規則的離散変数の集合 S_{ir} から規則的離散変数の集合 S_r への写像変換 Ψ (一種のスプライン関数)を施した上で分散的数値表現を用いるという方法も考えられるが、探索空間の歪によりエネルギーの最小点に変化するおそれがある。

4. 数値計算例

ここでは、ニューロンによる各種数値表現法を用いてニューロ・オブティマイザの挙動を調べる。Fig. 4に示す2変数の非線形関数の最小点を探索する問題を考える。ただし、各変数は0から20までの整数値をとるものとする。深い谷を渡る問題も難しいが⁶⁾、Fig. 4のように連続するなだらかな起伏を越える問題も難しい。この場合、規則的離散変数であることから分散的表現を用いることができる。3通りの温度パラメタ T_0 に対してエネルギーEの変化する様子をFig. 5に示す。 T_0 を高く設定すればニューロ・オブティマイザは真の最適解(5,5)を見つけ出している。次に、規則的離散変数に対して意図的に局所的表現および中間的表現を用いてみる。Fig. 6にエネルギーの変化する様子を示す。分散的表現の場合に比べると、状態が定常となるまでに時間を要している。Fig. 6の例では、中間的表現、局所的表現の両方とも真の最適解を見つけ出しているが、一般に局所的表現の場合には準最適解に止まることが多い。

不規則的離散変数に対して中間の数値表現が有効であるが、対象とする離散変数がベクトル変数の場合、変数間に大小関係の順序づけができないことがある。その場合には、中間的表現の数学的な根拠は消失し、局所的表現を用いなければならない。

5. おわりに

組合せ最適化問題に対する準最適化ツールとしてニューロ・オブティマイザの概念を提案した。また、0-1ニューロンによる離散変数の数値表現法がニューロ・オブティマイザの挙動に及ぼす影響について調べた。数値計算例を通して以下のことが明らかになった。(i) 準最適解へ到達するまでの計算量は、分散的数値表現を用いた場合に比較して、中間的表現、局所的表現では大きく増大する。(ii) 不規則的離散変数に対して中間の数値表現は有効である。

なお、大阪府立大学工学部学生 石橋貴嗣、高木信宏の両君からは数値計算および図表作成の協力を頂いたことを記し感謝致す。また、数値計算にはNEC PC9800システムを使用したことを付記する。

参 考 文 献

- 1) Hopfield, J.J. : Electronic Network for Collective Decision Based on Large Number of Connections between Signals, United States Patent, No. 4,660,166, 1987.
- 2) Takeda, M. and Goodman, J.W. : Neural Networks for Computation - Number Representations and Programming Complexity, Applied Optics, Vol. 25, pp.3033~3046, 1986.
- 3) 岸 光男 他：神経回路網モデルによる船体中央横断面の離散的最適設計，関西造船協会誌，第214号，pp. 11~17, 1990.
- 4) Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by Simulated Annealing, Science, 220, pp.671~680, 1983.
- 5) たとえば，平垣内正和 他：非線形最適化のための拡張Hopfieldモデル，電気関係学会関西支部連合大会予稿集，pp.337~338, 1988.
- 6) For example, Aarts, E. and Korst, J. : Simulated Annealing and Boltzmann Machines, John Wiley & Sons, 1989.
- 7) たとえば，増井裕也，松葉育雄：ニューロコンピューティングの基礎と実際（後編），Computer Today, 45, pp.67~72, 1991.
- 8) Matsuba, I.: Optimal Simulated-Annealing Method Based on Stochastic-Dynamic Programming, Physical Review A, 39, pp.2635~2642, 1989.

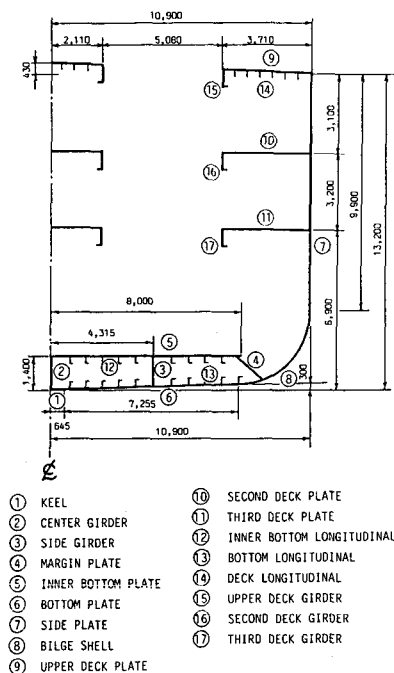


Fig. 1 Midship section of general cargo ship

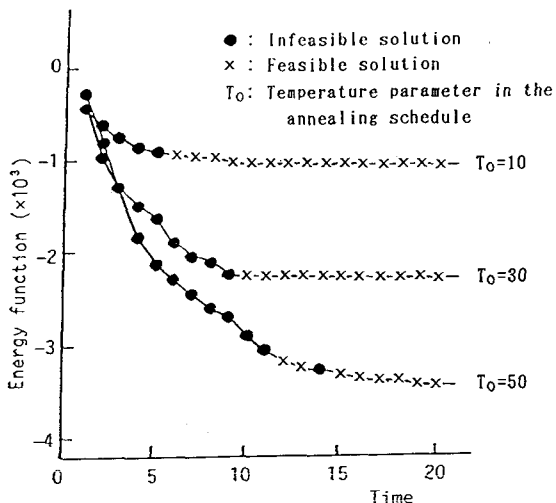


Fig. 2 Optimization process for midship section design

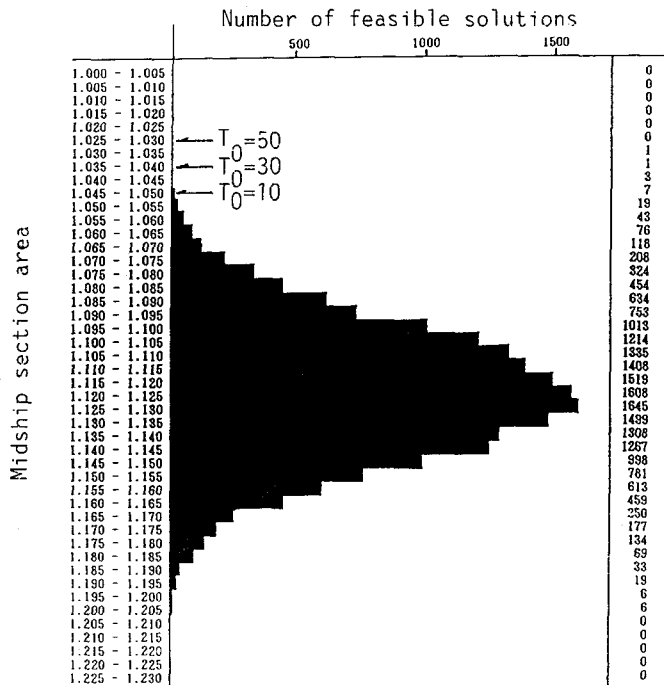


Fig. 3 Feasible solutions histogram for midship section design
(The number of samples is 20,000.)

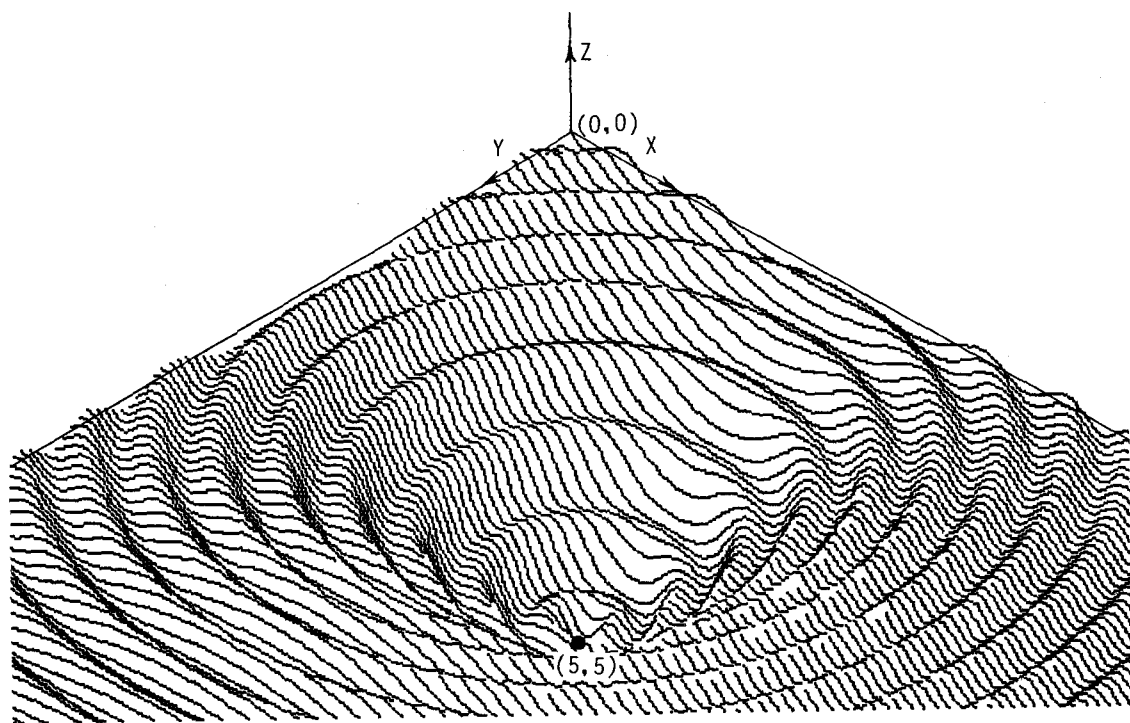


Fig. 4 Nonlinear function for discrete optimization

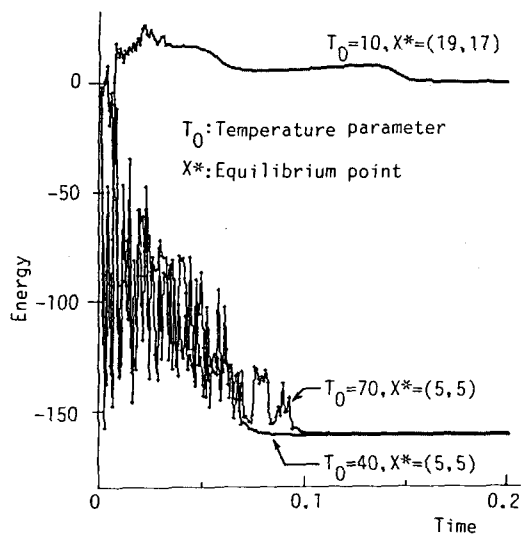


Fig. 5 Optimization process (for distributed representation)

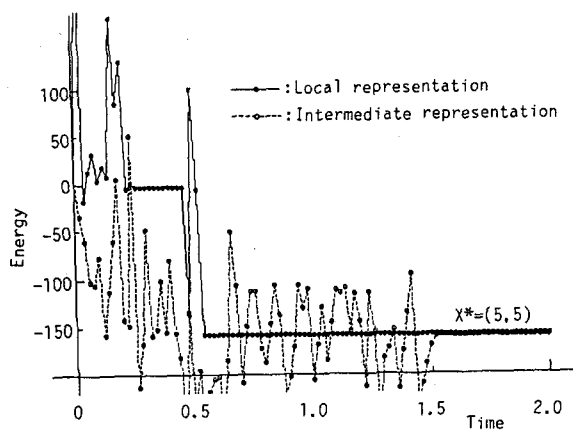


Fig. 6 Optimization process (for local/intermediate representation)