

## ボルツマンマシンに基づくマッチング機構

柴田 知博 井内 稔 山崎 清明

山梨大学

### 概要

データベースシステムに対する要求は、これまでの機械的なものから、次第に人間的なものに変わりつつある。また、扱うデータ形式についても、これまでの文字列などの単純なものに限らず、画像、音声も含めたマルチメディア化についての研究が盛んに行なわれている。本稿は、データベースにおける検索機構として、より人間的な処理が可能といわれるニューラルネットワークを用いたマッチング機構について考察したものである。

### 1 はじめに

高速なハードウェアの開発に支えられ、現代社会ではコンピュータを使うことのない生活は考えにくくなっている。このような社会の中で、さまざまな形でコンピュータが用いられている。ここで用いられているコンピュータはいわゆるノイマン型のコンピュータがほとんどである。このタイプの計算機は、プログラムされた処理を高速かつ正確に行なうことができるため、データベース検索など多種多様な用途に用いられている。

データベース検索システムにおいては、文字、数値、画像、音声など、さまざまな形のデータを扱えることが望ましい。

また利用する側としては、柔軟な検索ができることが望ましい。このような要求を満たすために、マルチメディアデータベースやオブジェクト指向データベースなど、近年盛んに研究が行なわれている。

一方ニューラルネットワークは、ノイマン型の計算機に比べて動作原理など不明な点が多く、しばらく地道な研究が行なわれてきたが、近年になって、より人間的な処理の実行が可能であることがわかってきた。このため、ニューラルネットワークの応用が盛んに研究されるようになってきた。

このような中で、検索する側の要求である「柔軟な検索」を行なうための機構について考察し、ニューラルネットワー

\*A new Matching Mechanism based on Boltzmann Machine.

クモデルの一つであるボルツマンマシンを用いて検証を行なった。

## 2 ニューラルネットワーク

ニューラルネットワークは、脳の情報処理をモデルにした情報処理機構である。この基本的な考え方は、1940年代に考案された。

脳は多数の神経細胞(ニューロン)によって構成されており、ニューロンが相互に複雑に結合することによって、さまざまな情報を処理している。ニューロンは、多入力1出力であり、一般にアナログ値を扱っている。他の細胞との接続の種類によって、信号を送る相手のニューロンの出力を増大させたり、減少させたりする。このことから脳は、多数のアナログプロセッサが相互に結合したものと考えることができる。

ニューラルネットワークは脳と同じような構造となっており、多数のユニットから構成されている。各ユニットは脳のニューロンに相当し、多入力1出力の処理機構である。各ユニットは他のユニットと複雑に接続しあっており、ユニットどうしの相互作用によって処理を行なう。

脳と同様に、ユニット間の結合には、あるユニットが興奮した時に周辺のユニットを抑制する結合(抑制性結合)と、結合しているユニットも興奮させる結合(興奮性結合)の2種類がある。ニューラルネットワークの学習は、この結合の強さ(結合荷重)を変化させることにより行なう。

ニューラルネットワークのモデルは、ユ

ニットの結合方法に従うと次の2つに分類することができる。[1]

**相互結合型** 情報の流れは双方向で、ユニッ

トは自分の出力を伝えた相手からも情報を受けとることができる。このため、自分の出力は別のユニットを経て再び自分への入力になることがある。

**階層型** 情報の流れは一方向で、入力層を構

成するユニットの入力が決まれば、順番に隣接する層を構成するユニットに対して情報が伝播する。

ノイマン型コンピュータはプログラムに忠実な処理しかできないのに対して、ニューロコンピュータは連想や推測など「あいまいな」情報を処理できる。ノイマン型で「あいまいな」情報処理を行なわせようとする、ある程度の性能は得られることがわかっているが、本質的に困難であることもわかってきている。逆に、ノイマン型コンピュータが得意とする論理的な処理はニューラルネットワークで実現するのは困難である。

### 2.1 ボルツマンマシン

ボルツマンマシンは、1984年に Hinton と Sejnowski により提案されたニューラルネットワークである。

ネットワークは  $\{0, 1\}$  の値を持つ2値ユニットで構成される。ユニットは可視ユニット群(visible units)と隠れユニット群(hidden units)の2種類があり、各ユニットは対称の結合( $w_{ij} = w_{ji}$ )を持つ。各ユニットの活性値は次式によって確率的に定まる。

$$p(u_i(t) = 1) = f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

$$x = \sum_j w_{ij} u_j + \theta_i \quad (2)$$

$u_i$ : ユニット  $i$  の出力値

$w_{ij}$ : ユニット  $i$  とユニット  $j$  の間の結合荷重

$\theta_i$ : ユニット  $i$  の閾値

$T$ : 温度 (確率関数の傾きを変えるパラメータ)

ボルツマンマシンは時間発展規則に従って動作させると、状態の出現確率がボルツマン分布で表されるような平衡状態に収束する。各状態のエネルギー値を決めるエネルギー関数は、ユニット間結合係数やユニットのしきい値などのパラメータなどによって決まるため、これらを調整することで望みの平衡分布を実現することができる。ボルツマンマシンには相互想起型と自己想起型とがあり、それぞれ次の特徴を持つ。

**相互想起型** 入力ユニット群、出力ユニット群、隠れユニット群の3種類のユニット群から構成される。学習は、入力ユニットに入力パターンを、出力ユニットには入力に対応する出力パターンを提示する。学習済みのネットワークを動作させると、入力ユニット群に示されたパターンから連想するパターン(学習時に提示された入力パターンの中で最も近いパターンに対応する出力パターン)を出力ユニットに出力する。

**自己想起型** 可視ユニット群、隠れユニット群の2種類のユニット群から構成される。相互想起型の特殊な例とみなすこ

とができ、入力と出力の区別がない。すべての入出力は可視ユニットに対して行なう。学習は、パターンを可視ユニット群に提示する。可視ユニットにパターンを提示し、学習済みのネットワークを動作させると、学習時に提示された最も近いパターンを可視ユニットに出力する。

### 3 マッチング

#### 3.1 マッチングの原理

マッチングには2種類ある。一つは、文字列のように記号化されている情報を対象とする。もう一つは、絵や写真のような記号化されていない情報についておこなうものである。

前者は、方法としてはいろいろなものが存在するが、いずれも各文字列を構成する文字がどの程度一致しているかをマッチングの基準としている。一方後者の場合は、画像の場合にはその画像を構成する部分の特徴を検出し、その特徴がどの程度一致しているかが基準となっている。

ここで人間が行なっているマッチングに注目すると、人間の場合はこのどちらのマッチングについても容易に行なっている。しかし、文字列のマッチングは明らかにコンピュータに行なわせている方法とは異なる。コンピュータの行なうマッチングは、文字を表すコードが数値的に等しいかどうかである。ところが、人間の場合は文字列を見たり読んでみたりして同じかどうかを判断

する。言い換えれば、文字列をコードとしてではなく画像や波形パターンとしてとらえ、パターンが似ているかどうかを判断することによりマッチングを行なっている。

人間はこのようなマッチングを行なうことにより、コンピュータが不得意である曖昧なマッチングをなんなくこなしているのである。そこで、曖昧さを扱うことが得意であるといわれる、ニューラルネットワークを用いてマッチングを行なうことを考える。

### 3.2 データ構造

ボルツマンマシンのパターン想起機能を利用するため、「物」をどのように表現すべきかを考える。通常、名称を忘れてしまった人が忘れてしまった「物」を相手に伝える時は、「物」の色や形などその特徴を列挙する。大まかな特徴だけではうまくいかない場合には、さらに細かな特徴を挙げる。つまり「物」の定義づけはより細かな「物」の特徴を挙げることにより行なわれている。

これに基づき、データ構造としてフレームを用いることとした。ユーザは各スロット値の入力を行ない、これをボルツマンマシンへと入力する。たとえ入力の一部が誤っていたり、わからなかったとしても、ボルツマンマシンのパターン想起により、正しい出力を得ることができる。

フレームを用いて表現した「物」をボルツマンマシンに入力しなければならないが、ボルツマンマシンは「0」か「1」の状態し

か扱うことができないため、何らかの形でコード化しなければならない。

### 3.3 データのコード化

ここでは、スロット値がとることのできる値は何種類かに限定しておき、それを値集合(表1)と呼ぶことにする。入力される特徴は必ずこの値集合に含まれる値であるという、制限をもうける。

図1に示すように、ユーザの入力値は必ず値集合に含まれていなければならない。この例の場合、入力項目「色」に対するユーザの入力は「緑」である。色についての値がセットされるスロットがスロット2であった場合、スロット2のコード変換テーブルを参照する。「緑」はコード「01111」に対応する。従って、ユーザの入力値「緑」はボルツマンマシンにセットする直前に「01111」というコードに変換される。後はこのコード1桁ずつボルツマンマシンの対応する可視ユニットにセットするだけである。

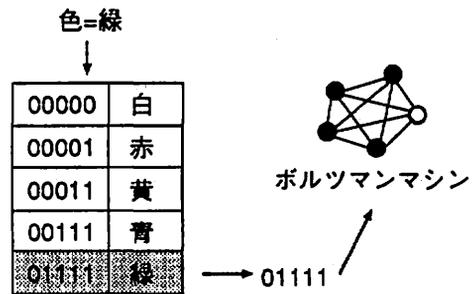


図1: コード変換

コードが「0」である時はユニットの活性化を「0」(ユニットが活性化していない)とし、

表 1: 値集合の例

特徴	値
色	赤、青、紫、黄、緑、橙、白、……
大きさ	-1cm, 1-3cm, 3-5cm, 5-7cm, 7-9cm, ……

コードが「1」の時はユニットは「1」（ユニットが活性している）を出力しているとする。

各スロットについて、値集合とそれに対応するコードの対を表したコードテーブルの一例を表 2 に示す。コード化の基準は、

- (人間が見て) 似ていると判断される特徴は、コード上でも同様に似ていると判断されなければならない。
- 一つのコード体系では定義できない特徴 (例えば色の場合は「白」や「黒」は、ほかのどの色とも「似ている」と判断できない。) を一つのコードテーブルに定義する場合は、「特別な」コードを用いる。

とし、「コードが似ている」とは「ハミング距離が近い」と、定義した。なぜなら、スロットを構成するユニットが全体として一つの意味を表しているため、一部のユニットが正しくない出力を行なっても全体として見れば正しいパターンとすることができるからである。

色を定義している表 2 では、白および黒は他のどの色とも近いものはないため、コード上では矛盾が生じている。しかし、白は淡い色のコードと、黒は濃い色のコードとそれぞれ近いように定義した。色の関係は

図 2 であるため、その他の色に関してはこれに従った。

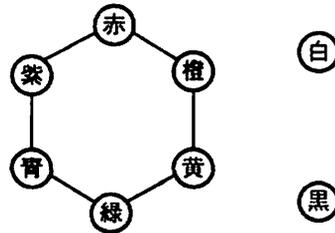


図 2: 色の関係

色のように 2 種類の基準があるためにコード化が困難なものについては、5 桁あるコードのうち上位 2 桁と下位 3 桁の 2 つの部分と見ることで、対処した。

#### 4 シミュレータの構成

今回の実験は、「昆虫データの問い合わせ」を行った。昆虫を表すために、データはフレームを用いて定義した。このフレームのスロット数は 5 である。各スロットには、それぞれ「名前」「色」「形」「大きさ」「羽」を表す値が定義可能である。

実験に用いたボルツマンマシンは、可視ユニットおよび隠れユニットの二つのユニット群からなるユニット総数 30 の自己想起型ボルツマンマシンである。可視ユニットは

表 2: コードテーブルの例

コード	データ	コード	データ	コード	データ	コード	データ
00000	白	01100	橙	11011	紺	10010	橙
00001	桃	01101	黄	11111	黒	10110	黄
00011	水	01111	緑	11110	深緑	10100	橙
00010	藤	01110	黄	11100	茶	10101	黄
00110	水	01010	紫	11101	黄土	10111	緑
00111	蓬	01000	赤	11001	紫	10011	青
00101	レモン	01001	紫	11000	深紅	10001	紫
00100	肌	01011	青	11010	茶	10000	赤

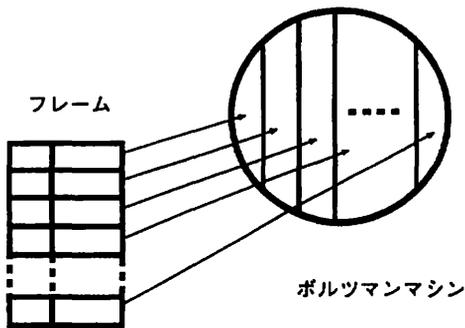


図 3: ボルツマンマシンとデータの関係

図 3に示すように入力するフレームのロット数に合わせて5つの領域に分割し、各ロット値はコード変換後、対応するユニット群にセットされる。

ボルツマンマシンのトポロジーは、すべてのユニットが互いに結合したものと、同じロットを構成するユニット群は完全に結合しておりロット間の結合は他の結合に比べて疎であるような結合のものの2種類を用意した。これは、扱うデータ構造が平坦なデータの羅列ではなく、フレームであることから、同一のロット値を表現する

ためのユニット群の結合は強く、ロット間の結合はそれに比べて弱くてもよいのではないか、という予測がたつためである。

## 5 実験

### 5.1 実験内容

実験用のデータの一部を表 3に示す。例えば、データ 1は「モンシロチョウという名前を持ち、色は白。胴体は厚みを持った細長をしており、体長は1～3cm。不透明な三角形の羽根を持つ」という意味である。

このデータを先に説明したボルツマンマシンに学習を行ない、以下の実験を行なった。

1. 学習済みのパターンを入力したときのマッチング成功率の測定
2. 学習済みのパターンに一部変更を加えてマッチング成功率を測定
  - 入力の一部を省略(パターン補間機能)

表 3: 実験用データの例

	名前	色	形状	大きさ	羽根
1	モンシロチョウ	白	厚細長	1-3cm	不透明三角
2	アゲハチョウ	黄	厚細長	3-5cm	不透明三角
3	カブトムシ	黒	厚楕円	4-6cm	楕円
4	クワガタムシ	黒	薄長円	3-5cm	楕円

- 入力に学習値に近い値を提示 (パターン修正機能)
- 入力に学習値と全く違う値を提示 (同上)

ただし、学習時に提示したパターンを入力した場合には、そのパターンを出力した場合に「成功」とし、パターン補間の実験においては、正しく補間した場合を「成功」とした。

また、それぞれの実験においては 1000 回のパターン提示の繰り返しを行ない、成功率として表示した。

## 5.2 実験結果

入力に対するボルツマンマシンの出力の、学習したパターンに対する照合率を図 4 および図 5 に示す。

図 4 は、スロット間の結合を減らした場合の成功率、図 5 は、すべてのユニット間に結合を持たせている場合である。いずれも実線で示しているグラフが正しいパターンを提示した時のマッチング成功率、点線は一部のスロットに変更を加えて提示したものである。また、横軸は「データの種類」、

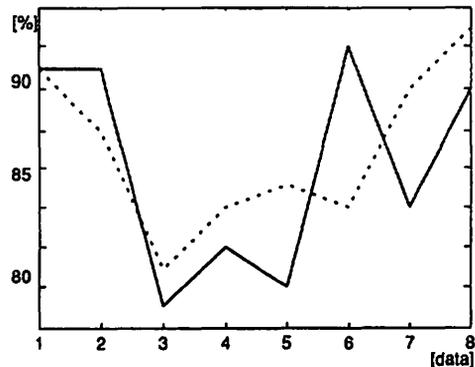


図 4: 結合を間引いている場合の学習パターンに対するマッチング成功率

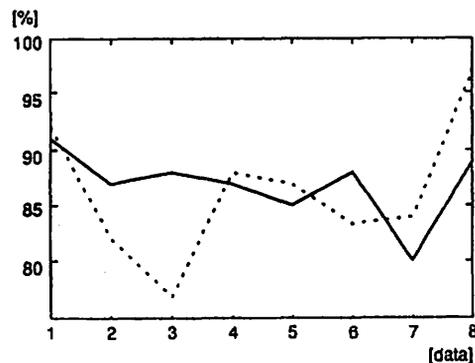


図 5: 完全に結合している場合の学習パターンに対するマッチング成功率

縦軸は「成功率(%)」である。

トポロジーを変更したものとそうではないものの違いはあまり出ていない。しかし、マッチングに失敗した(正確なパターンを出力しなかった)場合の出力を解析すると、完全結合のネットワークを用いた場合、誤って入力されているスロット値の影響のために正しい入力がなされているスロット値が変更されている例が多く見受けられた。したがって、完全結合のネットワークを用いるよりも、データ構造に合わせたトポロジーのネットワークを用いた方がよい結果を生むと考えることができる。

また、入力するデータが正確であればあるほど成功率が高く、入力データ中の負正確なデータの割合が多いと、成功率が極端に悪いという傾向がみられた。

## 6 まとめ

ボルツマンマシンなどのニューラルネットワークを用いたマッチング機構を実現するためには、本稿でも述べているように、コード化が非常に重要である。なぜなら、「似ているもの」は「似ているコード」として扱う必要があるからである。本研究ではコードが似ていることをハミング距離が近いと定義して行なった結果、ある程度の成果を得ることができた。しかし、これ以外の定義を行なった場合の検証を行っていないため、コード化については今後検証を行なう必要がある。

また、このようなマッチング機構を用いることで、文字列や数値の検索だけではな

く、画像や音声などを含めたデータに対しても検索できる可能性がある。

問題点として、多くのデータを扱うためには非常にたくさんのユニットが必要になる点が挙げられる。ボルツマンマシンを使う限り、ユニット数の増加は学習を困難にするため、何らかの対策が必要である。

## 参考文献

- [1] 中野: 「ニューロコンピュータ」, 技術評論社(1989).
- [2] D.E.Rumelhart, J.L.McClelland and the PDP Research Group(監訳・甘利): 「PDP モデル」, 産業図書(1989).
- [3] 「新情報処理技術に関する総合的調査研究 — 超並列・超分散処理WG報告書 —」日本情報処理開発協会(1989).
- [4] Marvin Minsky(安西祐一郎・訳): 「心の社会 “The Society of Mind”」, 産業図書(1990).
- [5] 上野, 石塚: 「知識の表現と利用」, オーム社(1987).
- [6] D.Touretzky, G.Hinton: “A Distributed Connectionist Production System”, COGNITIVE SCIENCE 12, pp.423-466(1988).