修士論文

神経網の高速シミュレーション手法

指導教官　　富田眞治 教授

京都大学大学院工学研究科
修士課程情報工学専攻

津邑 公暁

1998 年 2 月 19 日

神経網の高速シミュレーション手法

津邑 公暁

内容梗概

人間のように知的なものを作りたい — それは、多くの計算機科学者の夢であろう。我々の最終的な目標も、人間のような知能を作ることにある。では、「人間のように知的なもの」とは、どのようなものであろうか。

「人間のように知的なもの」を目指す研究といえば、人工知能研究がまず挙げられる。Deep Blue の勝利に象徴されるように、人工知能は現在までに大きな成功を収めてきた。その成果に異議を差し挟む余地はない。

しかし一方で、現在までの人工知能研究の延長線上に我々の望む「人間のように知的なもの」があるのかという素朴な疑念を禁じ得ないのもまた事実である。なぜなら現在の人工知能の産物は、特定の分野では人間をも凌ぐ能力を発揮する一方で、幼児でさえ可能な常識に関する問題や、犬や猫にでさえ可能な簡単な認識の問題をいまだ解くことができないでいるからである。これら人工知能の産物は、たしかに知的ではあるけれど、あまり人間には似ていないように思われる。

もちろん、人工知能のアプローチでは人間のような知能を獲得できないと断言することはできない。むしろ、原理的にはそれは可能であると考えられる。しかし、人工知能のアプローチは一見 目的に至る最短経路に見えて、実はそうではないのではないだろうか。

人工知能研究はその表象性において非常に具体的であり、明示的である。人間であるところのシステム設計者が自身の知識およびその処理について内省し、その結果得られた処理のイメージをモデル化することで、この知識を記号化を行っている。このようにして得られた知識表現は、その基本的構成が人間の内省に基づいた構造であるため、論理的推論などの人間の高位な思考を模倣するのには適したものとなる。

しかし我々の思考は、自身によって説明できるような過程を常に経るわけではなく、我々は時として理由を説明できないような判断を下したり、また判断を誤ったりする。このようなシステムが、人工知能と同様の表象性で実現するのは困難ではないだろうか。人工知能研究では人間の内省によって得られたアルゴリズムを用いるため、思考などの知的処理において人間にも説明が可能な

ことはそのアルゴリズムにとっても可能であるが、裏を返せば人間に説明ができないことはアルゴリズムにもできないのである。

我々は、人間のように知的なものの構築を目指すにあたり、人工知能的アプローチではなく人間の脳に答えを求めることにした。我々の最終的な目標は、「人間のように知的なものを作る」ことであって、必ずしも人間や脳そのものを作ることではない。しかし、我々は人間の脳よりも人間のように知的なものを知らない。したがって、我々の作るものが脳よりも知的ではない間は、脳を手本とすることは選択肢のひとつとして悪くはないと考える。

脳と人工知能との本質的な違いは、その表象性、とくに記憶の構造にあると考えている。脳の表象性を手本にしてシステムに記憶のメカニズムを導入するには、まず脳の記憶のメカニズムを知る必要もあるだろう。

そこで脳に答えを求めるにあたり、今回神経網のシミュレーションを高速に行うためのシミュレータを開発した。LSI の論理シミュレーションに用いられる離散時刻シミュレーションの手法を応用することで、計算量の低減を実現した。また、シナプス遅延を単位時間でなく実数時間で扱えるように、事象管理のためのタイムマップを binary tree で構成した。

作成したシミュレータを用い、神経細胞数 約 10000、各神経細胞あたりの接続線路数 約 100 という環境で測定を行った。全ての細胞に対して計算を行い単位時間の遅延しか扱えない従来の方法に対し、発火率 10% の状況では約 4.4 倍 の速度向上が得られた。また、本シミュレータは実数時間の遅延が扱えるにもかかわらず、神経細胞の発火率が 50% を超えない範囲においては従来の方法よりも高速にシミュレーションが行えることを示した。また、更なる高速化のために、同シミュレータを並列化するための手法についても考察を行った。

An Efficient Simulation Method for a Nerve Network

Tomoaki TSUMURA

Abstract

The dream of many computer engineer should be to develop a really intelligent system like human. The goal of my research work is to develop it.

Then, what's the "really intelligent system"?

The Artificial Intelligence (AI) research is one of the fields of research aiming at such systems. AI research has produced many results. It is symbolic of the success that Deep Blue the chess program beaten the human chess champion. I have no objection to good results of AI research.

On the other hand, I cannot help having some misgivings. Will present AI approaches reach such a really intelligent system at last? Within narrow limits, products of AI researches are exceed human beings all right. But they cannot able to answer common-sense-related queries even a child can, or they cannot recognize circumstances even a beast can. Of course, products of AI researches are intelligent. But they are quite different from us the human beings.

I cannot declare that present AI research approaches cannot reach a really intelligent system. If anything, I think present AI research approaches may be able to reach it finally. But, I feel that AI research approach would not the shortest way to our goal.

The representation in AI researches is very definite and very specific. System designer invents models of knowledge structure, and these models are based on his reflection on his own knowledge. These models are very effective for represent our logical consideration.

But, our daily behavior is not always well-grounded. Sometimes we cannot explain our own behavior, or sometimes we make a mistake. Will AI research products make a careless mistake? Knowledge models used in AI researches are very effective to represent our thinking process which we can recognize. It is because the models are based on designer's own reflection. But the reverse is also true. These models cannot represent our thinking

process which we cannot notice.

For accomplishing my purpose, I prefer human brain mechanism to AI approach. The goal of my research is not building an artificial brain but building a really intelligent system which comes close to human being. Our brains are most intelligent system as far as I know. Hence, human brain is relevant sample for my goal, I think.

The essential difference between brain and AI products should be the structure for their memory. When we would model after on brain memory structure, we have to know the memory mechanism of brain.

To know and to analyze brain function dynamics, I implemented a fast nerve network simulator. A parallel discrete event simulation technique cut down the quantity of computation. Nerve network model for my simulator is an extension of the pulse neuron model. With this simulator, synapse delays can be set up by not a unit time but a real number time. For this extension, I implemented the time map for events management by binary tree structure.

I evaluated the performance of this simulator with the nerve network model which have about 10,000 nerve cells and each nerve cell has about 100 synapses. Even in case of using a real number time delay, we obtain the following results.

- When firing nerve cell rate is **about 10%**, we get about **4.4 times** speed up.
- When firing nerve cell rate is **under 50%**, simulation speed exceeds usual simulation.

To make this simulator more faster, I weighed the parallelizing method for this simulator.

神経網の高速シミュレーション手法

目次

1 はじめに

「見わたすと、その檸檬の色彩はガチャガチャとした色の階調をひっそりと紡錘形の身体の中へ吸収してしまって、カーンと冴えかえっていた。」[1] このような比喩表現に出逢うことも文学の楽しみのひとつである。美しい比喩は我々に感銘を与えるのみならず、直截的表現よりも その描写をいきいきとしたものにし、より書き手のイメージを的確に伝えてくれる。

しかし我々は、どのようにしてこういった比喩表現を理解しているのだろうか。ここで言う理解とは、単に言葉の意味の置き換えではなく、上記のような書き手のイメージまで含めた理解を指している。例えば計算機にはこの比喩が理解できるのだろうか。逆に視点を変えて考えると、例えば比喩を理解できるようなシステムは、どうすれば構築できるだろうか。

ここで、比喩というのはあくまで一例であり、我々が問題にしたいのは人間の認識一般におけるその処理であるが、簡単のため、以下でも比喩を例にとって考えてみる。文学的に言うと比喩はレトリックの一種であり修辞的手段である。しかしそれはまた我々の日常生活における伝達をより効果的に行うための手段でもある。比喩を理解するためには、その背後にある意味や文脈を考慮にいれた推論が必要であり、そのためには言語の構造や意味だけではなくより一般的な知識が必要となる。計算機で実現しようとした場合、各機能をモジュール的に処理するにしても、最終的にそれらを統合的に扱えるような構造が必要であるだろう [2]。比喩を理解するにはまず、自然言語の理解能力が必要であることは言うまでもない。しかしこれだけでは不十分であり、比喩理解を実現するためにはその文脈や発話状況の認識能力およびある程度の常識なども備えている必要があると考えられる。

このようなシステムが実現できた状況を想像してみると、それはどのようなシステムであろうか。状況を理解し、常識を有し、言葉を理解する。それは十分に人間のような柔軟で知的な情報処理システムであると言えるだろう。この場合、比喩理解という処理は このシステムによって副次的に得られる能力の一にすぎない。このシステムが他の様々な能力を有するであろうことは容易に想像できる。

そのシステムはさまざまな知的システムに応用できるであろう。たとえば小さなものでは、意味や文脈を理解するためミスの少ない かな漢字変換 システ

ムであるかもしれない。たとえばそれは あなたの有能なアシスタントとなり得るかもしれないし、あなたと談笑のできるシステムとなり得るかもしれない。

我々の研究の目標は、このような知的で柔軟なシステムを構築することにある。

ここでひとつだけ断っておきたいのは、我々は 賢いかな漢字変換システム や比喩理解のシステムやアシスタントなどを作りたいのではないということである。それらの機能はあくまで副次的に得られるものであり、また副次的にしか得られないものであろうというのが、私の考えである。本研究の最終的な目標は、そのような副次的機能を満たしうるような統合的な認識システムの構築にある。

さて、このようなシステムを実現したいと考えるとき、まず思いつく方法としては人工知能的アプローチが挙げられる。人工知能は Deep Blue の勝利にも象徴されるように、現在までに大きな成功を収めてきた。しかし、人間のような知能を構成する手段として見た場合、多少の違和感を禁じえないのもまた事実である。

この違和感はどこからくるのだろうか。その理由のひとつは、人工知能の表象性にあると考える。人工知能における知識表現は、設計者であり人間である我々が自らの思考について内省した結果得られたイメージに基づいているため、論理的推論などの高位の思考過程を模倣するには適したモデルだと言える。しかしそれは非常に離散的であり、自分の思考過程と比較したときに あまりに似ていないことに気づかせられる。果たして従来の人工知能研究の延長線上に、我々の目指す 人間のように知的なシステム があるのだろうか。

数年来、このような漠然とした疑念を抱えていたが、一方で最近になって脳の研究が盛んに行われるようになってきている。1990 年 米国が「脳の 10 年 (Decade of the Brain)」宣言を行ったころから、脳関連の多くの国際科学連合組織が協力して、脳の研究が推進されてきている。大脳生理学の分野でも、PET(Positron Emission Tomography) や fMRI(functional brain Magnetic Resonance Imaging) といった無侵襲 (noninvasive) な脳の高次機能の計測手法が発達し、それらによる示唆的な実験事実から、脳の構造が従来我々が抱いていた脳のイメージとは多少異ったものであることが明らかになってきている。

脳のモデル化においても様々な研究がなされており、脳研究からのアプローチとしてもニューラルネットワークや、大脳生理学の研究結果を用いたものな

ど多岐にわたる。そんな中で我々のアプローチの特徴を一言で言うと、知識の単純な記号化を避け、構造や学習方法などの従来は明示的に与えられてきたものの正しいと思われる方法をシミュレーションを通じて発見することにある。

本論ではこのような目的のもと、前半において本研究のアプローチを他の研究との比較において明確にする。また後半では、本研究における脳の表象のモデルを考える際に必要となる神経網のシミュレーションを高速に行うため、今回作成した神経網のシミュレータについて述べる。以下、第 2章では本研究の背景として従来の人工知能研究および最近の動向について詳しく述べてその特徴を明らかにし、第 3章で、それを踏まえたうえで我々の研究のアプローチについて説明する。また第4章では関連研究について概観し、本研究との違いを示すことによって本研究のアプローチを明確にする。次に第 5章では、今回行った神経網のシミュレーションとその手法について述べ、評価結果を示す。最後に 6章でまとめを行う。

2 研究の背景

本章では、本研究と従来の研究との違いを明らかにするため、知的なシステムを目指す従来の研究の代表として人工知能研究をとりあげ、その特徴を述べる。

また、本研究の背景として、最近の生理学分野における研究成果やそれに基づく工学分野での動向について説明する。

2.1 人工知能研究

人間のように知的なものを目指す研究として誰もが先ず思いつくのは、人工知能研究であろう。

人工知能研究は大量の知識データに対して、高度な推論を的確に行うことを目指している。膨大な数値データに対して複雑な計算を高速に行うとことができるという意味においては、今日の計算機は頭脳に代わる働きを十分に実現している。そういった意味で現在の情報処理技術は、人間の能力をはるかに越えたものであるといえるかもしれない。

1956 年にダートマス大学で人工知能という概念が提唱されてから 40 年、Deep Blue の勝利に象徴されるように、人工知能研究は現在まで大きな成功を収めてきている。人工知能の応用としては、知識の活用を中心とする知識工学が提唱

され、その適用の１つとして専門家の経験的、技術的知識を電子計算機に移植し、活用しようとするエキスパートシステムが広く普及しつつある。

人工知能研究は、1950 年の Turing の論文 [3] に大きな影響を受けているとされている。この論文の主旨は、

- Turing Test の提案。
- 知能を実現するには、学習機構を用いる手段が有望である。
- 全てを学習で行う必要はなく明示的に形式論理を組み込むと効率がよい。

という 3 点 に要約できる。

その後の人工知能研究は、この最後の「明示的な形式論理の組み込み」という考えに基づき、

1. システムの設計者が、人間の知能とは何かを内省する。
2. それを計算機によって計算可能な形に定式化する。
3. それを計算機に与え、計算を実行させる。

というステップを踏むことで知能が実現できるという理念のもとに成り立ってきた。

以下にこの人工知能研究のアプローチの特徴を挙げ、それらについて考えてみる。

2.2　人工知能アプローチの特徴

人工知能のアプローチの特徴は、おおまかに言うと

- 「全てのことがらを記号化する」という表象性
- 「アルゴリズムは人間が書き与える」という明示性
- 「正しい解が存在し、それに適合するようにシステムを調整する」という目的性

の 3 つ にまとめられるだろう。この 3 つ の性質は互いに絡みあうものであり、完全に分離して考えることはできないのであるが、以下でそれぞれについて詳しく説明する。

2.2.1　表象性

人工知能研究は、人間も計算機もともに「記号」に対する処理を行っており、この能力こそが人間の知的行為の本質である、という前提に基づいている。ここでいう記号とは、人間が直接扱う文字や言語のみならず、計算機が扱う記号まで含めたものを指している。このような抽象度の高い記号は、一般には表象

と呼ばれる。

人工知能研究の表象性は、その記号処理的性質に代表されるだろう。3.1節でも詳しく述べるが、人工知能研究ではシステムの設計者であり人間である我々が、先ず自身の知識やそれに対する処理について内省し、その結果として得られたイメージをモデル化することによって表象を実現する手法がとられる。

つまり、人工知能研究において表象 (記号) として用いられるのは、人間である我々の意識にのぼるだけの高い抽象度をもった概念に対応するものであり、比較的メタな構造であると言える。例えば意味ネットワークなどにおいては、単語などの我々が思い描くことのできる概念が、ノードという構造にと対応づけられ実現されている (図 1)。

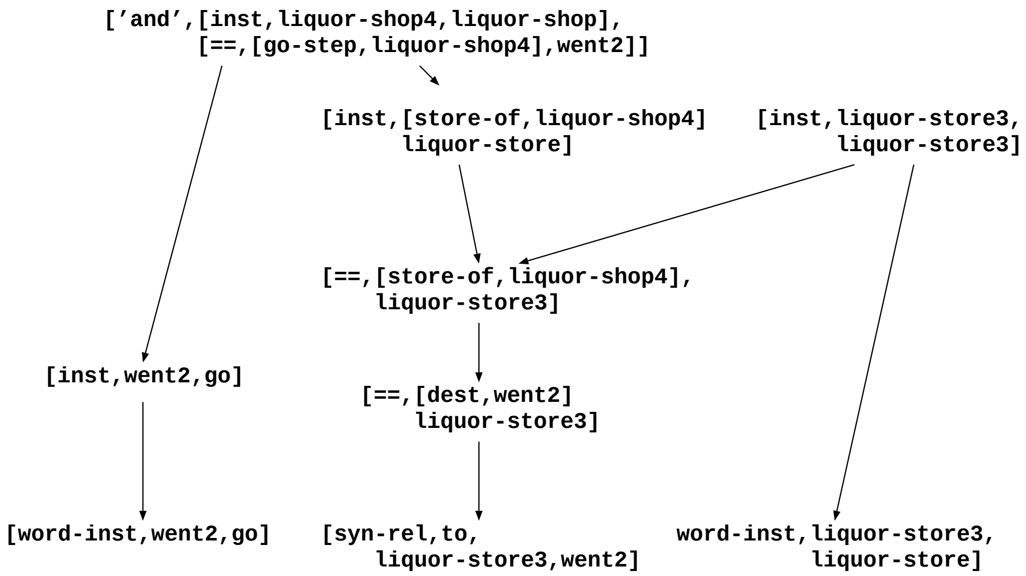

図 1: Bayesian Network モデルの例

このようにして構成された知識表現は、例えば論理的な推論など、人間の高位の思考過程を処理するのには非常に適していると言える。これは、我々が自身で行っている論理的思考をモデル化しているのであるから、当然といえば当然である。

基本的に人工知能研究は、このようにして構成された表象性の高い概念モデルに対し、手続き的操作を行うことによって知的行為をシステム上で実現しようという考えかたに基づいている。

2.2.2 明示性

前項のとおり、人工知能研究は基本的に人間の知能に対する内省に出発点があり、それに基づいて定式化を行うことで計算機で知能を扱うというアプローチをとる。

このようにして得られたシステムは設計者の内省に基づいているため、当然ながら知識モデルの記号化のレベルすなわち表象性が高く、理詰めによって行うような人間の高位の思考を表現したり模倣したりするのには非常に都合がよい。

例えば前項の例の意味ネットワークを構成するノードにしても、またプロダクションシステムにおける規則などにしても、我々のなかに それと一対一に対応する概念が存在していることは我々自身で理解できるため、このようなモデルは人間の知識の構造をある程度よく表しているといえる。

このようにして構成された知識モデルだけでなく、それに対する処理の方法も同様に設計者の内省によって与えられる。解りやすい例としては、GPS (general problem solver, 一般問題解決システム) において人間の思考過程を調べるのに用いられた 言語プロトコル (verbal protocol) 分析と呼ばれる手法が挙げられるだろう。

GPS は、初期の人工知能研究で知能を把える方法として用いられた方法で、

1. 目標と現状の差を検出
2. 差を縮める処理の検索
3. 検索により得られた処理の適用

という手続きを、目標が達成されるまで繰り返すものである。この GPS に、適用すべき処理の決定手法などの細かな点を具体的に実現したものがのちのプロダクションシステムに相当する。

この GPS に用いられた言語プロトコルという方法は、心理学で用いられる発語思考 (thinking aloud) と呼ばれるものと同様の方法である。それは人間の被験者に、問題を解いているあいだ「思考の内容を声に出して喋る」ことを要求することによって被験者の思考過程を追跡しようというものである。そしてそれを解析して得られた処理過程をシステムに模倣させようと考えたのである。

上の言語プロトコルは処理過程に関する例であるが、同様に人工知能研究では

- 知識の表現形式 (記号化方式、表象)
- 知識モデルに対する処理や操作手順
- 学習アルゴリズム

などの様々なものが設計者によって明示的に与えられることで、効率よく人間の思考に近い処理を実現しようとしている点に特徴がある。

最近では、知識獲得、類推、学習といった Turing への原点回帰的な動きもあり、必ずしも明示性が強いとは限らない研究も行われているが、それらのためのアルゴリズム自体は やはり人間が与えたものであり、その点ではあまり変わっていないと言えるだろう。

2.2.3 目的性

システムにとって、ある具体的な「進むべき方向」が決まっており、それに近づくようにシステムを導くことを指して、ここでは目的性と呼ぶことにする。

人工知能研究のアプローチにおける目的性を示す例としては、以下のようなものが挙げられるだろう。

- 学習アルゴリズムの適用
- 問題解決におけるサブゴールの設定

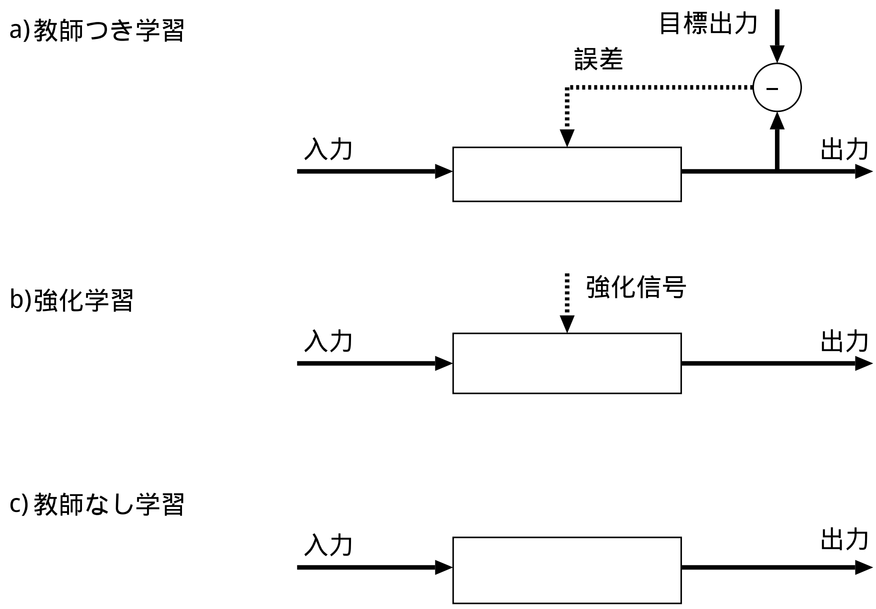

図 2: 学習アルゴリズムのパラダイム

人工知能における学習アルゴリズムは、その多くがニューラルネットワーク

などにも共通して用いられることが多いが、これらの学習アルゴリズムには大きくわけて以下の 3 つ のパラダイムが存在する (図 2)。

- 教師つき学習
- 教師なし学習
- 強化学習

教師つき学習 (supervised learning) では、個々の入力パターンに対して目標出力パターン (教師信号) が与えられ、実際の出力と目標出力との誤差をもとに学習が進められるもので、ニューラルネットワークの誤差逆伝播 (back propagation) 法などはこの顕著な例である。

その対極にある教師なし学習 (unsupervised learning) では、個々の入力に対する目標出力パターンというものは存在せず、入力信号の統計的分布を手がかりに、その中の特徴的変数を抽出する。

強化学習では、目標出力パターンは明示的には与えられず、強化信号 (reinforcement) あるいは報酬 (reward) と呼ばれる、出力がどれくらい良かったか悪かったかという評価にあたるスカラ値だけが与えられ、それに基づいてより良い評価につながる出力パターンが探索的に学習される。

一般にこれらの 3 つ の学習のフレームワークは、必ずしも排他的なものではなく、例えば全体としては教師つき学習の課題で、入力信号の前処理に教師なし学習を使ったり、あるいは強化学習のシステムの一部に教師つき学習の要素を用いるなどといったことも行われる。

人工知能研究では、設計者によってこのような学習アルゴリズムが適宜システムに与えられ、システムはそのアルゴリズムに沿って学習を行う。このように、学習における目的性は明確である。入力として与えられる、システムが解決すべき問題には、そのシステムが生成すべき応答の正解が存在し、それに適合するようにシステムの調整が行われる。

また、もうひとつの人工知能研究における目的性の例としては、問題解決を効率よく行うためによく用いられる、サブゴールの設定手段である。解決すべき問題は、設計者によってより小さな問題へと分割され、システムはそれらを順次こなしていくことでステップを踏みながら全体の問題解決に至る。これによって処理や学習が効率的に行われる。我々人間の思考も、このようなサブゴールを用いてステップを踏みながら順次 行われる場合もあるため、それなりに妥当な方法であると考えられる。

人工知能研究に限らず、システムにとってある機能を実現するという目的が存在している以上、このような目的性が存在するのは当然のことである。しかし人工知能研究では、その目標となる形に現システムを近づけるためのシステム自体の更新方法は、高いレベルの規則が設計者によって与えられるのが一般的であり、その与え方が明示的であるところに特徴がある。

2.2.4 まとめ

これまで見てきたように、人工知能研究には顕著な特徴が存在している。ここで、あらためてこれらの特徴をまとめてみると、以下のようになる。

表象性　人間も計算機も、行っているのは「記号処理」であるという考えのもと、知識を記号化して表現する。

明示性　知識の記号化方式や、それに対する処理、あるいは学習方法など、ほとんど全てを設計者である人間が書き下してシステムに与える。

目的性　常に正解が存在し、それに適合するようにシステムの変更がなされる。システムの変更のしかたは、やはり明示的に与えられる。

これらの特徴により構成されたシステムは、人間の高位における思考過程を表現するのは非常に適している。また、我々が自らの知識や思考について省みた場合に描くイメージとよく一致したモデル化でもあり、直観的に理解しやすいという利点もある。

しかし、果たして人工知能はこのアプローチで真に知的なものにたどり着くことができるのだろうか。

人工知能の成功は、前述した 3 つの特徴がうまくあてはまる分野に限定されているといえる。人工知能は、特定の分野の複雑な最適化問題などで人間を遥かに凌ぐ能力を発揮する一方で、犬や猫にでさえ可能な認識の問題をいまだ解くことができないでいる。

もちろん、人工知能のアプローチでは人間のような知能を獲得できないと断言することはできない。むしろ、原理的にはそれは可能であると考えられる。しかし、実際にそれを実現しようとしたとき、物理的にそれが可能であろうか。

J. Searle が 強い AI を批判するのに用いた有名な例に Chinese Room というのがある。ある閉鎖された部屋を考え、その部屋の中には、中国語による質問とその答えの精密で十分巨大な対応表 (辞書) が用意されているとする。また、この部屋の壁には一箇所、薄い紙が通るくらいの細い隙間が空いている。さて、ここで中国語の全く理解できない一人の人間を、この部屋の中に入れる。そう

して壁に空いた細い隙間を通じて、外から中国語で書かれた質問が部屋の中に挿入される。中の人間は、その中国語で書かれた文字列を、意味を理解しないまま巨大な辞書で引き、それに対する答えを紙に書いて同じ隙間から外へと出す。この部屋全体は外から見ると中国語を理解しているように見え、Turing Test に合格する。

この Chinese Room は、人間のように知的なシステムとしてよいだろう。つまり我々の目指すひとつの形ではある。しかし、人間の知識に相当すると考えられるその精密で十分巨大な辞書は、果たして用意することが可能なのだろうか。

時間をかければ用意できると考えても、それはあまりに膨大すぎて物理的に実現するのは可能な気がしない。それは、我々の描く辞書のイメージが、人工知能の知識表現におけるそれと同様に表象的に過ぎるからではないだろうか。我々が自身の知識について内省したときに得られるイメージはあまりに離散的すぎ、実際の知識の保存形式としては向かないのではないか。人間の脳という「生きた証拠」が存在するのであるから、そのような膨大な辞書も、うまくすれば我々の脳と変わらないサイズで実現できるはずなのである。しかし現在のような記号化の方法では、到底そのようなサイズで実現できそうな気がしない。

また、知識表現のみならずそのアルゴリズムにも同様のことが言える。人工知能で用いられるような高位の思考過程をうまく表現できる表象的なシステムにおいては、自分の思考過程が説明できるようなことに関してはうまくいくだろう。

しかし逆に考えると、人工知能の成功は人間の知的行為のうち意識的、論理的な部分を模倣したものに限定されていると言える。考えてみれば それは当然で、人工知能研究では人間の内省によって得られたアルゴリズムを用いるため、思考などの知的処理において人間にも説明が可能なことはそのアルゴリズムにとっても可能であるが、人間には説明ができないことはアルゴリズムにできないのである。

それでは果たして我々が目指す知能の中には、人間には説明できないことが含まれてはいないだろうか。その答えは否である。我々は自分で説明できないような思考過程を度々経るのであって、そのようなことも現在まで用いられてきたようなアルゴリズムで実現できるとは考えにくい。従って、もっと表象レベルの低い何かが必要であろうと考える。

2.3 最近の動向と脳研究

人工知能的アプローチに対し、このような漠然とした疑念を抱えていたが、一方で最近になって脳の研究が盛んに行われるようになってきた。

最近の生理学研究の進歩により、脳内の働きはかなり分かってきている。PET(Positron Emission Tomography) や fMRI(functional brain Magnetic Resonance Imaging) [4] といった、高次脳機能の無侵襲 (noninvasive) な計測技術が開発され、またこれらの分解能も技術によって上昇している。脳内のどれらの分野が記憶に関係しており、どういう順序で活性化するかなどといった、脳内の大局的な動きはおおまかにではあるが解明されつつある [5]。

また、個々のニューロンなどの細かい動きも、かなり分かってきている。LTP(long term potentiation)/LTD(long term depression) と呼ばれる、シナプス電位の長期形状変化 [6] が記憶の本質であることが分かってきており、それらはシナプス前膜 /後膜で生じる それぞれタンパク質 /グルタミン酸受容体のリン酸化が原因であることも知られるようになった。また、 LTP が大量に発生している状態では飽和した状態になり学習が行われないという実験結果なども得られている。

しかし、このように個々のニューロンレベルの細かい動きや、大脳においてどこでどういう機能が担われているかといった大ざっぱな機能局在が明らかになっていく一方で、ニューロンの群のような中間的なレベルで脳を見たときには、まだその動作が十分に理解されていないのが実情である。そしてまた、脳の情報処理を知ろうとしたときに、この中間的レベルが最も重要な点であることも事実である。個々のニューロンの動きが詳細に解明されたとしても、ミクロな計測の追求は、例えるなら得体の知れない装置を構成しているトランジスタを調べつづけるのと同様で、装置全体の処理に対する解明に繋がるとは限らないだろう。

大脳生理学分野における発見に伴い、計算機科学の分野でも新しい動きがでてきている。まず、ニューラルネットワークにおけるモデルの見直しである。

従来のニューラルネットワークの研究では、形式ニューロンモデルにおけるシグモイド関数の利用などを見ても分かるようにニューロンの平均発火頻度を基にした頻度コーディングが仮定されてきた。しかし、実際の神経細胞の間での信号伝達にはパルス情報が関与していると見られるため、このようなモデルではパルス情報が失われているぶん生体のニューロンと同等の表現能力も失っ

ている可能性があると考えられる。

これに対し、パルスの到達時刻や時間間隔の情報をも考慮したコーディングを仮定し、より生体のニューロンに近いニューロンモデルがパルス動作型ニューロンである。パルス動作型ニューロンモデルは、ニューロンの膜電位や閾値電位の減衰も考慮したモデルである。このモデルでは、ニューロン間の信号伝達においてはパルス頻度だけでなくパルスの到着時刻や時間間隔などの時間パターンによっても情報を表現することが可能であり、従来のニューロンモデルの限界を越えたネットワークモデルが構成できる可能性がある。

その他に脳の周辺研究として、脳型コンピュータ研究、脳型コンピューティングと呼ばれる分野も最近になって現れてきている。これらについても多くの研究があるが、通常のニューラルネットワークモデルの限界を越える新しいモデルの研究というのが、ひとつのオーソドックスな方向である。大脳生理学における脳の動きの観測結果をモデルにとりいれ、それを忠実に再現してシミュレーションを行うことで人間の頭脳と同等な知能を発現させようとするアプローチもある。大脳生理学などで得られた知見を基に脳のモデルを仮定して、工学的に構成し、その動作を観察することで脳の仕組みを機能的に解明しようとするアプローチである。

しかし人間の知能はメソ・スケールの物理現象であり、このような還元主義的とも言えるアプローチでは見えてこないのではないかと私は考えている。

3　研究のアプローチ

人工知能研究は、前章で述べたような 表象性、明示性、目的性という特徴を生かし、大きな成功をおさめた。しかし、果たして人工知能はこのアプローチで 真に知的なものにたどり着くことができるのであろうか。たしかに人工知能研究でも 知識の獲得、類推、学習といった 分野が盛んになるにつれ明示性は薄らいでいるとも言えるが、これらのためのアルゴリズム自体はやはり人間が与えたものであり、そういった意味ではあまり変化していないと考える。

人工知能の成功は、前述した 3 つの特徴がうまくあてはまる分野に限定されているといえる。人工知能は、特定の分野の複雑な最適化問題などで人間を遥かに凌ぐ能力を発揮する一方で、犬や猫にさえ可能な認識の問題をいまだに解くことができないでいる。

また、先に挙げた人工知能研究アプローチの 3 つ の特徴のうち表象性と明示

性は、一般的なディジタル電子計算機によって計算を行うことからの要請であり、あまり本質的ではない。一般的なディジタル電子計算機は、手続き的で、ディジタルで、ビット並としての記号を扱うのに長けているため、そのような方法をとっているだけである。

もちろん、人工知能のアプローチでは人間のような知能を獲得できないと断言することはできない。ただ我々は、このアンバランスに、茫漠たる疑念を禁じ得ずにいた。

一方で、近年の大脳生理学分野の研究成果は、脳が今まで思われてきたものとは若干異なるものであることを示しつつある。

我々は、このような状況を鑑みて、人間のように知的なものの構築を目指すにあたり、人工知能的アプローチではなく人間の脳に答えを求めることにした。我々の最終的な目標は、「人間のように知的なものを作る」ことであって、必ずしも人間や脳そのものを作ることではない。しかし、我々は人間の脳よりも人間のように知的なものを知らない。したがって、我々の作るものが脳よりも知的ではない間は、脳を手本とすることは選択肢のひとつとして悪くはないだろう。

本章では、研究目的を実現するためにこれから我々がとるべきアプローチとその理由について述べる。また類似研究との比較を行い、違いを述べることによって研究の立場を明らかにすることとする。

3.1　表象性

脳を手本とするにあたって、人工知能と脳はどこが違うかということを明らかにしておく必要がある。我々は、人工知能と脳の本質的な違いは表象性にあると考えている。以下、人工知能研究における表象性と、人間の脳における表象性について概観する。

3.1.1　人工知能研究の表象性

人工知能研究のアプローチによって知能が実現できる、という論拠は、簡単にまとめると以下のようになるだろう。

> 計算機も、人間の脳と同様に記号を記憶し処理する能力を持ち、この能力こそが人間の知的行為の源である。
>
> したがって、人間がどのように知識を記号化し記憶しているのかを解明し、記憶された知識と外界の事象との組合わせ方法を探索するプ

ログラムが与えられれば、人間と同じように問題を解決する計算機システムが構築できるはずである。

ただし、人間が扱う (目で見たり手で書いたりする) ことのできる記号は、計算機の扱う記号とは同じものではない。つまり人工知能すなわち計算機が扱うことのできる記号は より抽象的性質を持つものであって、これは一般に表象と呼ばれる。表象主義とは、人間の知覚内容や人間の持つ知識などがこの表象によって記述され、また推論などの思考は表象に対する手続き操作であるとみなす考え方である。

このような表象の操作が形式的に記述できると仮定する。すなわち表象の持つ意味や内容にかかわらず、その形式的な特性にのみ注目することで操作可能であると仮定すれば、表象の操作は計算ですますことができる。つまり人間の思考も計算機のプロセスも記号の形式的特性にのみ基づいた処理を行っていると考えることで、両者が機能的に同等であるとみなすのである。

たしかに脳は、記号を扱う能力に長けている。人間は自らの知識を言語などの記号として表現することができるし、また言語という記号なしには複雑な思考は不可能である。また、人工知能で模倣されるような、人間の脳の高位の機能ほど、記号に依存している割合が高いと言えるだろう。

3.1.2 脳の表象性

脳の表象性に関する有名な例に、 Grandmother Neuron 仮説がある。この仮説は、脳の AIT 野 (下側頭回前部) と呼ばれる部位においてヒトやサルの手に対して特異的に反応する「手細胞」が発見されたことに端を発する。

それまで形態の認識に関しては、せいぜい輪郭線の線分の傾きなどに反応する細胞が存在することくらいしか分かっていなかったが、「手」という極めて具体的な形態を認識しているかに見える細胞が発見された。形態の認識処理の流れにおいて、 AIT 野 に処理が移る直前までの部位では線分のような単純な刺戟に反応する細胞しか見つかっていないのに AIT 野 で突然 複雑な形態に反応する細胞が見つかったことで、

> AIT 野 には様々な種類の形態認識細胞が多数存在し、それで視覚認識が行われる

という説が起こったのである。この説では、「おばあさんをおばあさんと認識できるのは、それ専用のおばあさん認識細胞があるからだ」とするため、Grandmother Neuron 仮説という名で呼ばれる。

この仮説によると、AIT 野ではひとつの神経細胞がひとつの概念を表すような構造になっていることになり、これは人工知能研究における知識表現の表象性と非常によく似ている。

しかし、この仮説は脳内で知識が記号として格納されていることを、直接には支持していない。また、この仮説が正しいとすると認識対象の数だけ認識細胞が必要であることになってしまい、数に限りのある脳細胞でそれをカバーするのは無理があるだろうと思われる。

一方で、最近の大脳生理学研究の成果によって、脳の表象性が単純な記号化とは異なるものであることを想像するに有効であろうと思われる成果がいくつかあがっている。以下にその例のいくつかを示すことにする。

例えば前項で挙げた Grandmother Neuron 仮説については実際には Grandmother Neuron のようなものが存在するわけではなく、認識対象がより抽象化された図形特徴に分解されて、細胞の反応が生じることが分かってきている。しかもひとつの細胞と、その細胞が反応する図形特徴は一対一の厳密な関係で結びついてるわけではなく、ある程度の対応の幅が存在することも確認されている。言い換えると、複数の神経細胞が複数の特徴の認識を分担して行っているといえる。

また、大脳皮質はコラム構造をなしていることが知られているが、このひとつのコラム (約 10 万個の神経細胞から構成される柱状の構造) の中には、特徴の似た図形や近い図形に反応する細胞が比較的近い位置に存在するが、隣接する他のカラム内には全く異なる図形に反応するような細胞が並んでいることなども確認されてきている。つまり、コラム単位でもおおまかな認識の分担を行っていると考えられる。

その他には、対符号化ニューロンというものも発見されている。ある刺戟 (例えば視覚刺戟) を与え、一定の遅延時間を置いて同じ刺戟を与えたときの反応の見ることで記憶の観察を行おうとする遅延反応試験と呼ばれる実験がある。これにおいて、通常の神経細胞は最初の刺戟時と 2 回目の刺戟時に反応するが、ふたつの刺戟の間じゅう発火しつづけるような細胞が確認されている。しかし、この記憶を保持しているかに思われるニューロンはどんな刺戟にも反応するというわけではなく、いくつかの特定の刺戟にしか反応しない。それ以外の刺戟には他の特定の細胞が反応することになる。

このようにニューロンがある特定の刺戟と特別の結びつきをもってそれを記

憶するようになることはチューニングと呼ばれ、これ自体はそれほど新しい発見ではない。しかし、このようなニューロンが複数の対象刺戟に対して反応する場合、その刺戟どうしは連想されやすい関係にあることが分かってきている。つまり複数の刺戟に対してチューニングされたニューロンは、それらの刺戟の間の連想に寄与しているということである。このようなニューロンは対符号化ニューロンと呼ばれる。

このように見てくると、人間の脳における情報の記号化は単純なものではないことが分かる。大脳生理学分野における これらの事実の発見は、脳の記号化が 少なくとも計算機上の人工知能のそれのようには表象的でないことをものがたっている。言いかえれば、必ずしも ひとつの事象をひとつの記号に対応させるような単純な記号化を行ってはいないと言えるだろう。

3.2 記憶

脳の情報処理において鍵となるのは、その表象化の方法すなわち記憶のメカニズムにあると想定できる。システムに記憶のメカニズムを導入するに際し、脳の表象性を手本とするには、まず脳の記憶のメカニズムをよく知る必要があるだろう。よってここでは、まだ明らかになっていない点が多くある分野ではあるが、脳における記憶の特徴のひとつである 短期記憶 / 長期記憶 の別について現在 解明されていることについて説明し、脳内における記憶の保存形式について考えることにする。

3.2.1 短期記憶と長期記憶

脳を構成するシナプスには、絶えず変化する性質があり、これは一般に可塑性と呼ばれる。

軸索や樹状突起の位置は誕生時に確立されて変化しないが、神経細胞間の連結は、シナプス結合部において絶えず変化している。神経細胞と樹状突起のシナプス結合部には隙間があり、電気信号のインパルスが前者から後者にじかに乗り移ることはない。シナプスにおいて情報を伝える手段として使われているのは、神経伝達物質と呼ばれる化学物質である。シナプスの上流から放出された伝達物質が下流に存在する受容体と呼ばれるチャンネルに到達することで、パルスが伝えられる。

このようにシナプスにおける情報伝達は、神経伝達物質を介した化学的な過程であるが、このシナプスを通じた情報伝達効率が変化することをシナプス可塑性

という。この原因としては、受容体の感度の変化、化学過程の効率の変化、受容体の数や形態の変化、あるいはシナプスそのものの増加などが考えられる。

可塑性のあるシナプスによって神経細胞を繋いでゆくと、たとえば記憶のような機能を持った神経回路を作ることができると考えられる。脳の高次機能はすべて、そのような神経回路によって担われていると考えられる。この可塑性を端的に表したのが Hebb 則であり、ニューラルネットワークの学習などにおいて広く用いられている方法である。 Hebb 則については、 4.2節で詳しく説明する。

最近の研究では、シナプス可塑性、特に新たなシナプスの生成やシナプスの形態変化などに遺伝子の関与が大きいことが分かってきている。

記憶はニューロンのネットワークの中につくりだされ、保持されている。ひとつのニューロンは別々の違ったニューロンからシナプスを介して入力を受けており、複雑な電位が絶えず生みだされている。そうしてつくられた電位変化を、また多くの別のニューロンにシナプスを介して送りだす。その電位の形やスピードや持続時間は、ひとつひとつのシナプスの効力の強さや形成されたシナプスの数、その位置、形などによって異なっている。記憶は、こうしてつくられる電位変化の形状が一定期間保持されることに対応するらしい。このような形状変化は実際にみつかっており、 LTP (long term potentiation) および LTD (long term depression) と呼ばれている [6]。

シナプスにおいて入力線維に反復的な興奮が起こると、シナプス前膜から化学伝達物質であるグルタミン酸が放出される。グルタミン酸受容体には AMPA 型 と呼ばれるものと NMDA 型 と呼ばれるものがあり、ともにイオンチャンネルも兼ねている。 AMPA 型 は Na^+ を通し、 NMDA 型 は Ca^{2+} を通す。また、 AMPA 型 は膜電位に関わらず常に同じ動作をするが、 NMDA 型 は脱分極が起こらないとチャンネルが開かないことが知られている。つまりニューロンが著しく興奮した状態でないと、働かないということになる。

シナプスの上流側から放出されたグルタミン酸は、シナプス後膜上の AMPA 型受容体に結合してそのイオンチャンネルを開き、 Na^+ が細胞内に流れこんでシナプス後膜を脱分極させる。 NMDA 型受容体のイオンチャンネルが開き、細胞内の Ca^{2+} 濃度が上昇する。 Ca^{2+} 濃度の上昇により、キナーゼ (タンパク質のリン酸化を行う酵素) が刺激され、受容体のリン酸化が起こる。リン酸化された AMPA 型受容体は感度が上昇し、ここにシナプス後膜における LTP が

発現する。

このとき、シナプス後膜における Ca^{2+} 濃度の上昇は、同時に逆行性伝達物質と呼ばれる物質を生成し、それがシナプス前膜側に拡散する。これによってシナプス前膜においてもリン酸化が発生し、シナプス前膜における LTP が発現する。このように、シナプス前部と後部の細胞が連携を行いながら LTP は発生することが知られている (図 3)。

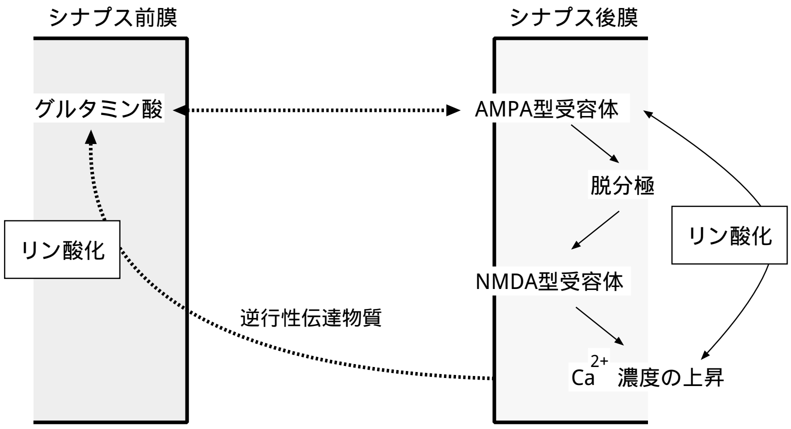

図 3: シナプス前膜、後膜の連携による LTP の発生の概要

人間の記憶には、大きくわけて 陳述記憶 と呼ばれるものと手続き記憶と呼ばれるものがあるとされている。前者は通常われわれが「記憶」という言葉からイメージする記憶であり、後者は自転車の乗りかたや泳ぎかたを覚えるというときの、いわゆる体で覚える記憶を指す。そしてこの 2 種類の記憶は、さまざまな記憶障害の症例から、脳の別の場所に宿るお互い独立した仕組によって保持されていると考えられている[1)]。

これらの記憶は、その内容がなかり恒久的に保存される性質のもので、脳の中から繰り返し取り出されて利用される類のものである。これは長期にわたって残る記憶、長期記憶 と呼ばれる。これとは対照的に、ほんの短期間だけしか残らない記憶も存在する。手続き記憶は反復トレーニングを前提とするため短期記憶というものを想定することはできない。よってこの短期 /長期という 2 種の記憶の区分は、陳述記憶にのみあてはめられる。

1) この陳述記憶と手続き記憶の例のように、脳内の独立した部位の破壊によってそれぞれの記憶が失われるような場合、その二者において「トウバーの二重乖離原則」が成立すると言う。

従来は あらゆる記憶に対して先ずこの短期記憶が形成され、そのあと一定期間を経て改めて長期記憶へと変換されるのだと考えられてきたが、短期記憶と長期記憶の間にも二重乖離原則が成り立つことがわかってきている。すなわち、必ずしも短期記憶の段階を経ずに形成される長期記憶もあるということである。

脳において記憶に大きく関与している部位として、側頭葉の海馬がよく知られている。

海馬に脳障害が生じた例において、その障害が起こる以前の記憶は残っているにもかかわらず、新しいことが覚えられなくなることが知られている。このことから、海馬は短期記憶を長期記憶に変換する役割を担っていると考えられる。

また、以降のできごとが記憶できなくなる前行性健忘のみならず、海馬に障害が起きた時点から遡ってしばらくの間の記憶が失われる逆行性健忘の症状も見られる。このことから、短期記憶は海馬において一定の保管期間を経て、長期記憶へと変換されると考えられており、人間では症例からその期間は約 2 年であろうと推測されている。

3.2.2 脳における記憶の表象

脳の記憶について考えるとき、考慮すべき点として先ずひとつは、脳のフォールトトレラントな構造が挙げられる。先に挙げた Grandmother Neuron 仮説の例で考えると、この仮説が正しいとするとき、 Grandmother Neuron が死んだ場合は「おばあさん」が認識できなくなることが考えられる。しかし実際にそのような筈はなく、人間は脳細胞がひとつ死んだところで記憶量が 2^{-1} 倍になるわけではない。

これは脳におけるデータ構造が単に冗長性を持っているためだとは考えにくい。それよりもむしろ複数の事象に対する記憶を複数の神経細胞が分担して引き受けていると考えるのが自然であろう。そうすると 神経細胞がひとつ消えたときにも ある記憶の脱落が起こるのではなく、複数の記憶が薄くなるだけですむ。

次に、情報落ちと情報圧縮の問題がある。人間の記憶量は脳細胞の数から想像される記憶量をはるかに凌いでいるように思われる。これと同様な記憶量をコンピュータ上に実現することは現在のところ不可能だとも感じる。しかしそれは記号による表象性の高い記憶を想定しているからであって、それとは別の

記憶のメカニズムが存在するのであれば状況は変わってくる。

人間の記憶とコンピュータにおける記憶のもうひとつの大きな違いのひとつは、その曖昧性にある。人間の記憶はコンピュータのそれのように厳密な正確さを必ずしも有していない。これは、記憶時に情報落ちのある圧縮に似た変換が行われているためと考えることもできる。

このように、脳における記憶のメカニズムは非常に複雑であり、また連続的である。現在用いられている知識表現のように、高い表象性をもつ記号とは別のメカニズムが存在しているように思われる。我々が、従来の単純な記号化とは異なった別の表象性を目指すのはこういった理由からである。従来の人工知能研究のように、記号化されたシステムに対する処理方法を模索するのではなく、記号化の方法自体が我々の研究対象であると言える。

我々の脳は、ニューロンと呼ばれる神経細胞から構成されており、それをモデル化したものがパターン認識などに用いられるニューラルネットワークである。ニューラルネットワークは我々の脳を模倣したモデルであるため、脳の機能を実現する際には比較的よいモデルであると想像できるが、実際のモデル構築の際には必ずしもニューラルネットワークに拘ることはない。このようなデータ構造を有機的に実現する方法としてはニューラルネットワークが最適であったとしても、コンピュータ上で実現する際に人間と同じ構造が最適であるとは限らないからである。

上述したような曖昧性や耐故障性を考えたとき、ホログラフィの理論やスペクトル拡散などもモデル化の際の参考になるのではないかと思われる。

3.3 アルゴリズムの獲得

問題処理のためのアルゴリズムを学習により獲得できるシステムの実現は容易ではない。一般的には決定不能な問題であり、アルゴリズム不在のシステムを考えることになってしまう。しかし、これはチューリングマシン上での話 すなわち離散系を前提とした場合である。アルゴリズム獲得のための連続系の特徴を持つ学習アルゴリズムはまだ あまりよいものが発見されておらず研究の途中である。

例えばニューラルネットワークにはホップフィールドネットワークと呼ばれるモデルがある。これは最適化問題のためのニューラルネットワークであり、

ニューロンの動作は

$$\begin{aligned}\frac{dU_i}{dt} &= -\frac{U_i}{\tau} - \frac{\partial E}{\partial V_i} \\ U_i(t+\Delta t) &= U_i(t) + \frac{dU_i}{dt}\cdot\Delta t \\ V_i = f(U_i) &= \frac{1}{2}(\tanh(\lambda\cdot U_i)+1)\end{aligned}$$

U_i ニューロン i の総入力

V_i ニューロン i の出力

τ 時定数

λ 入出力関数 f のゲイン

E エネルギー関数

で記述される。このネットワークモデルは、ビットパターンの獲得を連続系の特徴を生かして高速の求める近似解法を与えるものである。しかし解決の対象となる問題からビットパターン問題への変換は人間が行うのが通例であり、この変換を自動的に行うアルゴリズムの問題は解決されていない。

また、 cellular automaton を学習的に獲得させるような試みも、連続系の特徴が陽に導入されているわけではないし、その複雑さは甚大である。写像を確率化したシステム、すなわち出力される記号が決定論的には定まらず、それぞれの記号が出力される確率だけが規定されているシステムである隠れマルコフモデル (hidden Markov model: HMM) は、連続系がうまくオートマトンに導入された例と見ることもできるが、適応可能な問題には限界がある。また、このような確率を用いる方法には、その確率自体を求めるのが非常に困難であるという問題点があるうえ、確率ネットワークで表現できることは、コストに基づく仮説推論の論理体系でも表現できる程度のことでしかないため、望むような柔軟性は得られないと考える。

3.4 研究の方法

我々は、人工知能的アプローチではなく人間の脳に答えを求めることにした。我々の最終的な目標は「人間のように知的なものを作る」ことであって、必ずしも人間や脳そのものを作ることではない。しかし、我々は現時点において、人間の脳よりも人間のように知的なものを知らない。したがって我々の作るものが脳よりも知的ではない間は、脳を手本とすることは悪くない選択肢である

と考える。

現在の脳の研究の周辺は、以下のふたつに大きくわけることができる。

- 神経生理学を中心とした解析的研究
- 計算機科学を中心とした構成的研究

前者は生体を解剖あるいは計測し、神経細胞や神経系の仕組を物理的に解明するアプローチであり、後者は前者によって得られた知見を基に脳モデルを仮定して工学的に構成し、その動作を観察することで脳の仕組を機能的に解明しようとするアプローチである。

大脳生理学に代表される解析的研究の分野では、無侵襲な高次脳機能計測の手法が発達してきている。fMRI や PET, MEG などを代表とするこれらの手法において時間的および空間的な分解能が向上してきたことによって、これまで困難であった前頭葉の機能計測や生活環境下での脳の活動状態など、脳内のおおまかな動きの観測が可能になり、人間の高次機能や感情に伴う脳の動作がより詳細に調べられようとしている。

しかしこれは言ってみれば還元主義的なアプローチである。将来、仮に現在よりも技術が発達し、高次脳機能計測手法の分解能が極限まで向上したとする。そうすれば脳内の信号の動きなどが逐一観察できるであろう。しかし、果たしてそれによって脳の機能が解明され得るであろうか。知能はメソ・スケールな物理現象であり、このような還元主義的な手法では解明できないと私は考える。

もう一方の構成的研究では、計算機上に脳のモデルを構成しようとするとき、その規模の大きさが問題になる。脳モデルは例えば人間の脳が 10^{10} 規模の神経細胞の行っている処理としての並列アルゴリズムを対象にしているのであるから、その計算量は膨大である。この対応として特定の脳モデルを対象にした専用のハードウェアあるいは種々の脳モデルを対象にした汎用の計算機が開発され、その方式で実現できる機能を確認するとともにより広い対象・機能を実現しようとする努力が続けられてきている。

また最近では、脳型コンピュータ研究などに代表されるように、従来の人工知能研究とは違ったアプローチで人間の頭脳に迫ろうとする研究が多くなされている。しかし、それらは一概に脳の情報処理の解明 /実現に注目したものであり、記憶に関しては従来同様、記号処理的・表象的 な構造を処理系の外部に想定しているものが殆どである。

我々は 3.2節で述べたように、記憶にこそが脳の情報処理の本質であり、最

も重要とすべき部分であると考える。そこで我々のアプローチは、まず人間の記憶のメカニズムを探るため、短期・中期・長期の記憶がシミュレートされるような脳のミニチュアを作ることである。シミュレーションによって得られた情報からその記憶における性質を解析し、システム構成の際の参考にすることが現在の目標である。

先の人工知能研究の特徴と照らしあわせて我々の研究のアプローチをまとめると以下のようになる。

表象性　認識過程において我々の意識にのぼらない現象をも包含しうるモデルを構成する必要があると考えるため、ことがらを単純に記号化する方法は避ける。記号化の方法自体が研究対象である。

明示性　単純な記号化を行わないことにも関連するが、そのアルゴリズムの与えかたが分かっていないため、それが分からないあいだは明示的にアルゴリズムを与えることは極力避ける。

目的性　人間の処理と同様の処理が行えることを目標とするため、ある程度の正解は存在する。そのためその正解に適合するようにシステムを調整する必要があるが、やはりシステムの調整のしかたが分かっていないため、それが分からないあいだは生理学的に裏付けのある方法以外は極力使わない。

4　関連研究

本章では、本研究と同様の目標をもつ研究分野をとりあげて説明し、それらの特徴と本研究との違いについて述べることで本研究の立場を明確にする。

4.1　人工生命

人工生命研究の目的は、生命における現象を計算機内にシステムとして実現し、生命の本質に迫ることにある。その生命現象には、もちろん知能も含まれており、知能を生命現象のひとつと把えてその実現を目指している。

現在の人工生命分野の研究ににおいて最も強い主張は、

> 「科学のフロンティアには複雑系の科学があり、それを生命系で用いれば創発 (emergence) が生じる。そしてこの創発こそが人工生命の本質である」

というものである。

創発を説明するのによく用いられる例としては、蟻の隊列がある。個々の蟻

は隊列を作ろうと意図して行動しているわけでもなく、また隊列を作ることを指揮する個体がいるわけでもない。個々の蟻はフェロモンの分泌とそれに向かう走向性という単純な行動をしているだけであるが、そお単純な行動の結果として隊列という合目的的に見える現象が発生する。この例のように、ある階層に存在する属性だけでは説明できない属性が、より上位の階層で生じてくるという概念が創発と呼ばれるものである。創発システムとは、このような性質を利用して、実装の対象となる階層をより低いところに置き、そうしたシステムを動作させた結果として生じてくる属性を、本来目的としていた効果として利用しようとするアプローチであるとも言える。

人工生命研究では 基本的には第一原因を下部に求め、明示的なモデルの生成によるトップダウンアプローチは避ける。人間である設計者によって書き下ろされたモデルや方程式は、その設計者自身の偏見や独断に陥ることが多く、また動物にも行える簡単な環境認識など、明示的に処理方法を与えることができない問題も多いからである。

ボトムアップなアプローチを経て創発によって構成される生命は、設計者である人間が予めその構成を明示的に与えたわけではないため予測不能であり、また目的を持たない。目的を与えることは生命の構成の過程において「(設計者の考える) 進むべき方向」という目標を与えることになり、人工生命研究の柱である非明示性に反することになるからである。

以上まとめると、

表象性　設計者による記述は行わず、それらが創発によってシステムから発生することを期待する。

明示性　人工生命研究では、一切の明示性を避ける。第一原因最下部に求め、ボトムアップなアプローチを経て創発によって生命および知能を含む生命現象が実現されることを期待する。

目的性　目的を与えることは、生命の構成過程において設計者がシステムの手をひくことになり、本来の方向とは異なる方向に進んでしまう可能性がある。明示性が排除されているシステムは予測不能であり、無目的であることを目指す。

このように、人工生命研究の特徴としては、非明示性および無目的性が挙げられる。人間による介入を極力避けることで自然に生命を発生させようとするのである。換言すると、この点では明示的理解を放棄しているとも言える。我々

のアプローチと大きく異なるのはこの点である。

4.2 ニューラルネットワーク

ニューラルネットワークは、パターン認識などによく用いられる認識のためのモデルである。

人間の神経細胞を模した形式ニューロンと呼ばれる単純なモデルを組み合わせてネットワークを構成したものがニューラルネットワークである。形式ニューロンは多数の入力線とひとつの出力線を持つ構成をしており、各入力線に対する重み (シナプス結合荷重) および細胞自身の興奮の閾値を変数として持っている。形式ニューロンは入力信号が与えられると、各入力線に設定された重みを考慮した入力荷重和を計算し、その値と自身の閾値とを比較するステップ関数によって出力を決定する。入力を x_i $(i = 1, 2, ..., n)$、出力を y とすると、

$$y = \begin{cases} 1 & (u \geq 0 \text{ のとき}) \\ 0 & (u < 0 \text{ のとき}) \end{cases} \tag{1}$$

但し、 u は入力の荷重総和

$$u = \sum_{i=1}^{n} w_i x_i - \theta \tag{2}$$

となる。

ネットワークモデルは、層状の変換ネットワークがよく用いられる。単純なニューロンモデルでは、上述の式ように 実際の脳同様 u の値は $0/1$ のディジタル値をとる。しかし、階層状のニューラルネットワークで用いる場合には u に連続値をとらせるモデルもよく用いられる。その理由は、ネットワークモデルの空間識別能力にある。

この際、値 u に 飽和特性関数を作用させるのであるが、この飽和特性関数としてよく用いられるものがシグモイド関数である。

シグモイド関数を用いるときは、式 (1) の代わりに、

$$y = \frac{1}{1 + \exp(-u)} \tag{3}$$

を用いる。

式 (3) の u に 0 を代入した式は、 $X = \{x_1, x_2, ..., x_n\}$ の張るパターン空間

上では超平面を表す式となる。[1]つまり言い換えると、その超平面でパターン空間をカテゴリ分割する能力をこのモデルは持つことになる。

階層構造をもつニューラルネットワークでは、入力層 node の荷重を変更することによって中間層 node の出力であるシグモイド関数の勾配が制御される。更に、中間層と出力層間の荷重を変更することによって、中間層 node のシグモイド関数が 拡大 / 縮小 / 反転 される。これらによって、十分な数の中間層を持つ階層ニューラルネットワークは、任意の非線型関数を生成することができるのである。

このようにニューラルネットワーク研究は、神経細胞のモデルである形式ニューロンから始めている点で当初は我々同様、知的システムの構築を目指していたとも考えられるが、現在のニューラルネットワーク研究はモデルであるニューラルネットワークのパターン認識能力 (空間識別能力) に注目し、それをパターン認識の効率的手段として用いているにすぎない。神経網を扱っているという点以外に、我々との研究の共通点はないと言える。

ニューラルネットワークをコンピュータのアーキテクチャとして見たときには、ホップフィールドネットは非同期回路を、リカレントネットは同期回路を、シグモイド関数は離散系を越える連続系を表す能力を持つという意味で、形式的には現在のフォンノイマン型コンピュータと同等あるいはそれ以上のものを構成可能である。

しかし、これはあくまで形式的なニューラルネットワークの表現能力の話であり、脳のようにプログラムあるいはアルゴリズムを学習によって自ら獲得するようなシステムとはかけ離れている。プログラムやアルゴリズムを考えるのは依然プログラマである人間である。つまり、現状では処理可能な問題は基本的に「パターン認識」であり、考えるシステムの実現には至っていない。

我々の望むようなシステムは、過去の研究の延長線上にはないと考えられる。従来の学習方式の加速器ではなく、またシンボリックな AI でもない何か別の新しいシステムが必要である [7, 8]。

そのためにはやはりモデルの構築から始めなければならないだろう。しかし、モデルの能力を極力制限しないようにすることに十分留意しなければならない。簡略化され能力が制限されたモデルを用いてシミュレーションを行うと、その

---

[1] 結合荷重 w_i と閾値の重み θ によって傾きおよび位置が決定。

モデルに特化した特質しか見えてこず、またその情報によりモデルの最適化を行ってしまうため、ますます specific なモデルとなってしまう。最初の僅かな誤差[1)]がモデルの改良を重ねるごとに増幅し、最後に得られるものは当初望んでいたものとは異なるエキスパートシステムになってしまうからである。

また、表象性という面でニューラルネットワークを見てみると、層状ニューラルネットワークの出力層はまさしく前述の Grandmother Neuron であり表象的である。また中間層では、階層構造のニューラルネットワークでは上述のように非線形識別能力を左右する情報が中間層にシナプス荷重として保存されているわけであるから、中間層の荷重の分布がこのシステムの記憶であると言えるが、それに関する解析的研究は殆ど行われていない。また、当該ニューラルネットワークで処理される問題は普通特定されているため、中間層における記憶が明らかになったとしてもそれから記憶の解明に関して意味を持つ情報が得られるとも限らないと思われる。

また、ニューラルネットワーク研究にも目的性は存在する。神経細胞モデルにおいて、さまざまな学習モデルが存在するが、それらの殆んどが前提としているのが Hebb の学習法則 (Hebb 仮説) [9] と呼ばれるものである。

Hebb 則　神経細胞 A と神経細胞 B との間のシナプスは、 A の発火が B の発火に続いて生じたときにはその伝達効率が強化され、そうでないときには弱められる。それによって、神経網上に何らかの学習がもたらされる。

$\rho_{i.j}$ を i および j にある 2 つの神経細胞をつなぐシナプス結合強度、 A_i と A_j をそれぞれシナプス前およびシナプス後のアクティビティとするとき、上記の Hebb 則 によるシナプス結合強度の時間変化は以下の微分方程式で表される。

$$\frac{d\rho_{i,j}}{dt} = A_i A_j$$

ただしこれはシナプス結合の増加だけを記述しているため、シナプス結合は無限に強化され発散してしまう。これを回避するため、従来のモデルでは減衰項や拘束条件を付加している。

これを一般化した学習則として、学習信号 γ は、入力信号 $\boldsymbol{x}$、シナプス荷重 $\boldsymbol{w}$、教師信号 y によって定まる信号 $\gamma = \gamma(\boldsymbol{x}, \boldsymbol{w}, y)$ であり、シナプス荷重 w_i は、そのシナプスへの入力信号の強さ x_i とその細胞の学習信号 γ との積 γx_i

[1)] モデルの簡略化などにより生じる、モデルとモデル対象との違い

に比例して増加すると同時にわずかに減衰する。 Hebb 仮説は、神経細胞が興奮したときのみシナプス荷重の強化が起こるというものであるから、学習信号 γ としてその細胞の出力 z を採用することに相当する。

Hebb 則に基づく学習法では、神経細胞は自分が興奮したときにのみ学習が行われる。すなわち、出力値が $0/1$ の 2 値 の場合、学習信号を γ、細胞の閾値を h とすると細胞の出力を用い、

$$\gamma = z = 1(\boldsymbol{w}^T \boldsymbol{x} - h)$$

となる。

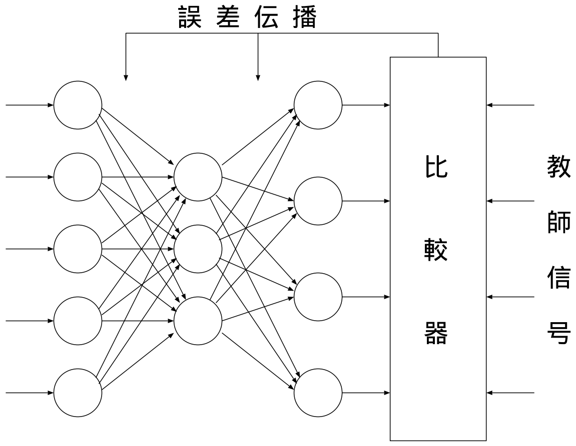

図 4: 誤差逆伝播

この Hebb 則 に基づいてネットワークを学習させるための様々なアルゴリズムが考えだされたが、層状ニューラルネットワークでよく用いられるのは、誤差逆伝播 (back propagation) と呼ばれるものである。層状ニューラルネットワークにおける学習とは、入力に対して望ましい出力を提示し (教師つき学習: supervised learning)、ネットワークが望ましい変換をするように結合荷重を修正することであると言える。ただしこれらの学習アルゴリズムはニューラルネットワークの目的性から発生したものであり、必ずしも生理学的な裏付けに基づいたものではない。つまり実際の生体の神経細胞においてそのような学習に相当する現象が発見されているわけではない。

また、 Hebb 則は出力がなされているまさにその時点での入力情報にしか効かない。つまり入力の時間的変化を学習に反映させることが困難である。何らかの形で時間的情報を学習に反映させることが必要であると考えられるが、そ

れがモデルを特定の方向に制限してしまっては人工知能研究のアプローチと全く同じ問題にぶつかってしまう。

また、たとえば脳と全く同じものを計算機内に構築することが実現できたとしても、それを学習させるには人間が生きているくらいの時間も必要であろう。

以上まとめると、

表象性　層状のネットワークでは、出力層が極めて表象的な役割を果たしている。学習後のニューラルネットワークにおいては、結合荷重がある種の表象であるとも考えられるが、そのネットワークが学習した特定のパターン認識問題についての表象でしかなく、またそれに対する解析的研究はほとんどなされていない。

明示性　学習を行わせるためには学習アルゴリズムを明示的に与えてやる必要があるため、明示性は避けられない。Hebb 則などの単純な規則だけ与えた場合は時間情報などが損失され、ネットワークの表現能力を制限してしまうため、より抽象レベルの高い学習則を与えるのが普通である。

目的性　教師つき学習、強化学習ともに期待値が存在し、それに出力および強化信号を近づけるように荷重の変更を行うという意味で、明確な目的性が存在している。

最近は形式ニューロンモデルはパルスの時間情報を反映しないという点が指摘され、より複雑で時系列情報を考慮したパルス動作型ニューロンモデルを用いた研究もなされているが、やはり記憶にあたるであろう荷重の分布に対する解析的研究はほとんどなされていない状況である。また、ニューラルネットワークの対象とする問題を制限していることもあって、荷重を解析しても問題に特有の性質は認められるかもしれないが、記憶一般についての性質を見極めることは、現在の方法では困難であろう。

なお、本稿におけるシミュレーションでは このパルス動作型ニューロンモデルをもとにしたニューロンモデルを用いているため、パルス動作型ニューロンモデルについては あとで詳しく述べる。

4.3　脳型コンピュータ

最近盛んになっている研究として、脳型情報処理あるいは脳型コンピュータと呼ばれる研究分野がある。この研究分野では一般に、大脳生理学分野の研究成果によって少しづつではあるが明らかになりつつある神経細胞の動作を忠実

に追いかけ再現することから、人間の頭脳と同様の機能を実現させようとする還元主義的なアプローチをとる。

従来から、神経生理学はじめ脳の研究において、知覚、運動、認知、記憶、学習、意識などといった、脳の高次機能が脳のどういうメカニズムによって担われているかという問題がある。

最近の技術の進歩や研究成果により、小脳に関しては、神経回路網の構造や機能がほぼ分かってきている。しかし肝心の大脳に関しては、

- 神経回路網がどういう構造になっているのか、またどういう機能をもっているのか。
- どういう情報が流れてどういう処理がなされているのか。

このような問題は、まだほとんど何も解明されていないといってよい。現在分かってきているのは、大脳のどの部位においてどのような機能が担われているのかといった機能局在くらいである。

現在の脳型コンピュータ研究は、ハードウェアで脳を実現しようとするアプローチをとっているものが多いが、そのほとんどはニューラルネットワークモデルのハードウェア化である。つまりニューラルネットワークの特徴であるパターン認識能力を利用し、高速化およびそれに伴う認識率の向上を目的としている。

しかし、中には単なるニューラルネットワークの高速化ではなく、ハードウェアによるシミュレーションによって現在のニューラルネットワークでは実現できない能力をもった新しいニューラルネットワークモデルをひきだそうとしているものもある。しかしやはり明示的にモデルの構成や学習アルゴリズムを与えざるを得ないため、従来のニューラルネットワークシミュレーションとあまり変わるところのない状況である。

5 シミュレータ

本章では、神経網の動作シミュレーションをソフトウェアで行うことの必要性について述べたあと、今回作成したシミュレータの構成およびその処理方式について考察を行う。

また、更なる高速化の方法として、シミュレーションの並列化手法および Sun UltraSPARC の Visual Instruction Set の利用について述べる。

5.1 神経細胞のモデル

今回のシミュレーションが前提とするモデルについて述べる。それに先だって、従来の単純なニューロンモデルや、最近よく用いられるパルス動作型ニューロンモデルについて説明しておく。

5.1.1 従来のニューロンモデル

従来のニューラルネットワーク研究においてはニューロンの発火頻度を考慮したコーディングに基づくニューロンモデルが利用されてきている。

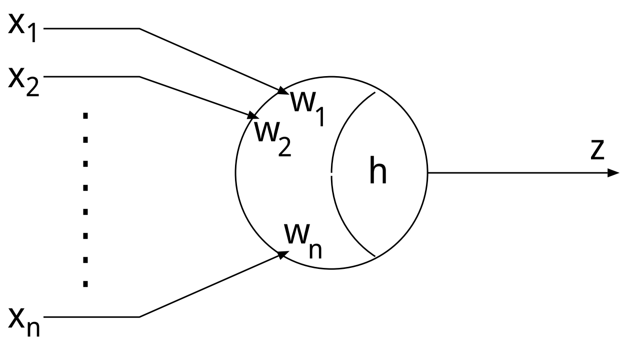

図 5: 単純なニューロンモデル

最も単純かつ よく用いられるモデルはMcCulloch & Pitts による形式ニューロンモデル (図 5) である。このモデルでは神経細胞は n 個 の神経細胞からそれぞれ $x_1, x_2, ..., x_n$ の強さの入力信号を受けとり、その膜電位 u および 出力 z は

$$\begin{aligned} u &= \sum_{i=1}^{n} w_i x_i \\ z &= f[u-h] \end{aligned}$$

で表される。ここで、 h は興奮の閾値である。また、 w_i はシナプスの結合効率を表す量で結合荷重と呼ばれるものであり、神経細胞の i 番目の軸索に単位入力が来たときに変化する膜電位の量を表す。興奮性細胞からのシナプスに対しては $w_i > 0$ であり、抑制性細胞からのシナプスに対しては $w_i < 0$ の値をとる。

McCulloch & Pitts が提案した当初のモデルでは、 f として単位ステップ関数が用いられていた。この場合、実際の神経細胞と同様に出力値は $0, 1$ のディジタル値をとる。しかし、一般には f はシグモイド関数のような単調増加関数

や単調非減少関数が用いられ、この場合ニューロンの出力は、0 から 1 のアナログ値をとり、そのニューロンの発火の平均頻度を表現する。

このようなニューロンモデルの集合であるニューラルネットワークモデルにおいて、階層構造を持つネットワーク (層状ネットワーク) は、そのパターンクラスに対する帰納的推論能力および汎化能力により、パターン認識やロボット制御に多く用いられてきた。

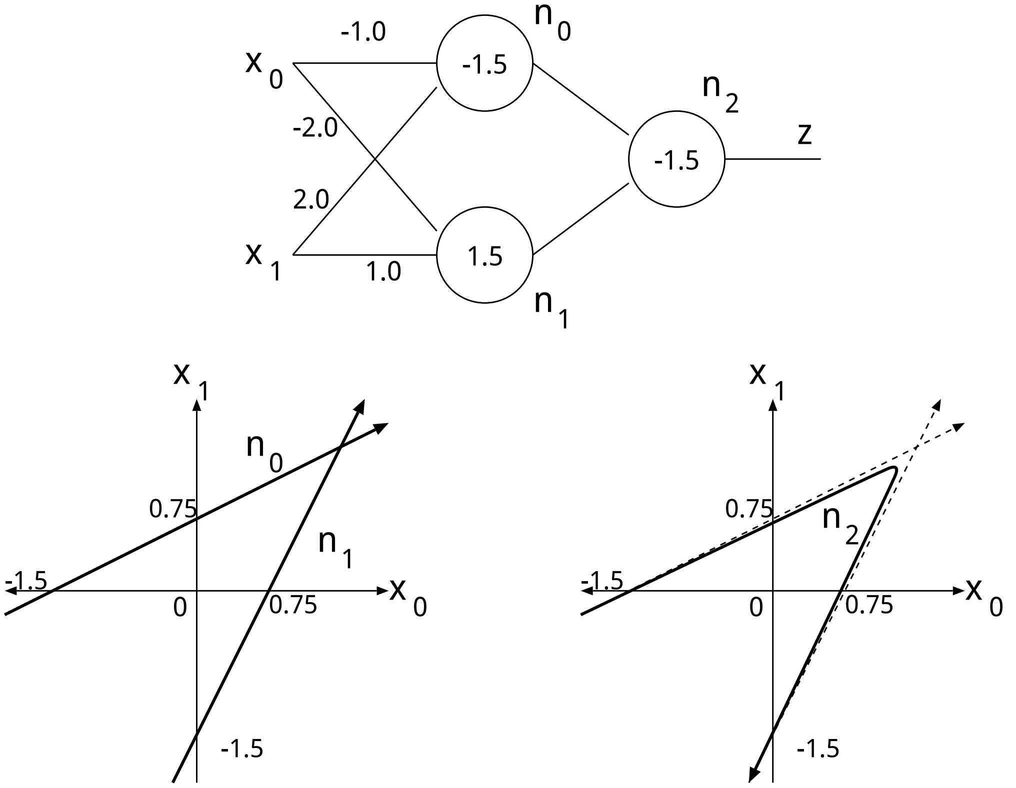

図 6: 層状ニューラルネットワークによる非線形識別

このような層状のネットワークでは、入力数だけの軸を持つ次元空間において、その中間層が超平面を構成し、それによりその次元空間を分割 (カテゴリ認識) するような線形の認識能力を持つ。よって、これらを出力層で統合することによって、層状のニューラルネットワークは非線形分離が行える。

図 6 のニューラルネットワークの例を考えてみる。ここで、ノード内の数値は閾値を表し、エッジに付してある数値は結合荷重を表している。このネットワークにおいて中間層のノード $n_0,\ n_1$ は、それぞれ図 6左下にあるような直

線を x_0x_1 平面内に構成する。この場合、これらの中間層の出力を統合する出力層のノード n_2 は、図 6右下の実線で描かれている境界を構成することができ、非線形の識別能力を持つことが分かる。

さらにこれらのネットワークの能力を引きだすため、入力にシナプスの結合荷重を乗じたものの和である入力総和にシグモイド関数 (図 7) などの飽和特性を作用させせることによって、非線形の識別能力を実現する方法も用いられる。

階層ニューラルネットワークでは、入力層のノードの荷重を変更することによって中間層ノードの出力であるシグモイド関数の勾配が制御される。更に、中間層と出力層間の荷重を変更することによって、中間層ノードのシグモイド関数が 拡大 / 縮小 / 反転 される。これらによって、十分な数の中間層を持つ階層ニューラルネットワークは、任意の非線型関数を生成することができる。

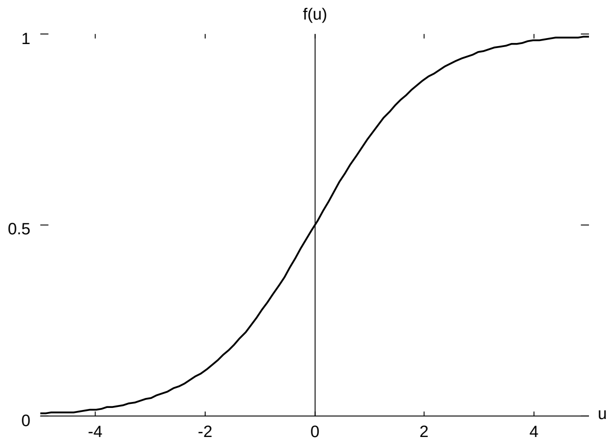

図 7: シグモイド関数の飽和特性

よく用いられるこのようなモデルでは、細胞の出力はシグモイド関数で表される細胞発火の平均頻度によってのみ決定されるため解析的な評価が行えるのが利点である。しかし、実際の脳内の神経細胞における信号伝達において観測されているパルス情報などは考慮されていない。つまり細胞間を伝わるパルスの到着時刻や到着間隔といった時間的情報がこのモデルでは失われてしまい、そのことがモデル全体の表現能力に限界を与えてしまっている可能性がある。

5.1.2 本シミュレータが前提とするモデル

これに対し、パルス到着の時刻や間隔の情報をコーディングに取り入れた、実際のニューロンにより近いモデルにパルス動作型ニューロンモデル [10] があ

る。

時刻 $t \geq t_0$ におけるニューロンの膜電位を $P(t)$、閾値電位を $H(t)$ とした場合、それぞれは以下のような式で与えられるとする。ここで、各変数/定数の添字 E, I はそれぞれ、興奮性、抑制性を表している。

$$
\begin{aligned}
P(t) &= P_\infty + (P_0 - P_\infty)e^{-(t-t_{I_{ji}})/T_p} \\
&\quad + \sum_{j=1}^{K_E} A_{E_j} \sum_{i=1}^{n_j} e^{-(t-t_{E_{ji}})/T_p} - \sum_{j=1}^{K_I} A_{I_j} \sum_{i=1}^{m_j} e^{-(t-t_{I_{ji}})/T_p} \\
H(t) &= H_\infty + (H_0 - H_\infty)e^{-(t-t_0)/T_h}
\end{aligned}
$$

P_0, H_0	初期リセット膜電位、初期リセット閾値電位
P_∞, H_∞	静止膜電位、静止閾値電位
A_{E_j}, A_{I_j}	j 番目のシナプスにおける単一パルスあたりの細胞膜電位の変化量
K_E, K_I	興奮、抑制 それぞれの入力線路数
n_j, m_j	j 番目の入力線路への入力パルス数
$t_{E_{ji}}, t_{I_{ji}}$	興奮、抑制パルスの印加時刻
t_0	パルス出力後、絶対不応期 τ を経た時刻
T_p, T_h	膜電位、閾値電位の時定数

膜電位 $P(t)$ はパルスが到着しない間は減衰を続けるが、パルスが到着するとその電位を急激に上昇させ、そののちまた減衰を始める。これに対し、閾値電位 $H(t)$ は単に減衰を続け、静止閾値電位 H_∞ に達するとそこで一定する。

時刻 t において膜電位 $P(t)$ の値が閾値電位 $H(t)$ の値を上回ると、その神経細胞は発火する。この際、 $P(t), H(t)$ は初期値 P_0, H_0 にリセットされる。これをそれぞれ初期リセット膜電位、初期リセット閾値電位 と呼ぶ。(図 8)

今回のシミュレータで前提とするモデルは、上記のパルス動作型ニューロンモデルを基として、さらに神経細胞の「慣れ」を考慮したモデルを用いている。

5.2 ソフトウェアシミュレーションの必要性

前章で述べたように、多少目的は異なるものの、大規模な神経回路網の高速シミュレーションが必要とされてきている。

従来のようなニューロンモデルでは、その構成および飽和特性の単純さから解析的に評価を行うことが可能であった。しかし、細胞の出力を細胞の発火の

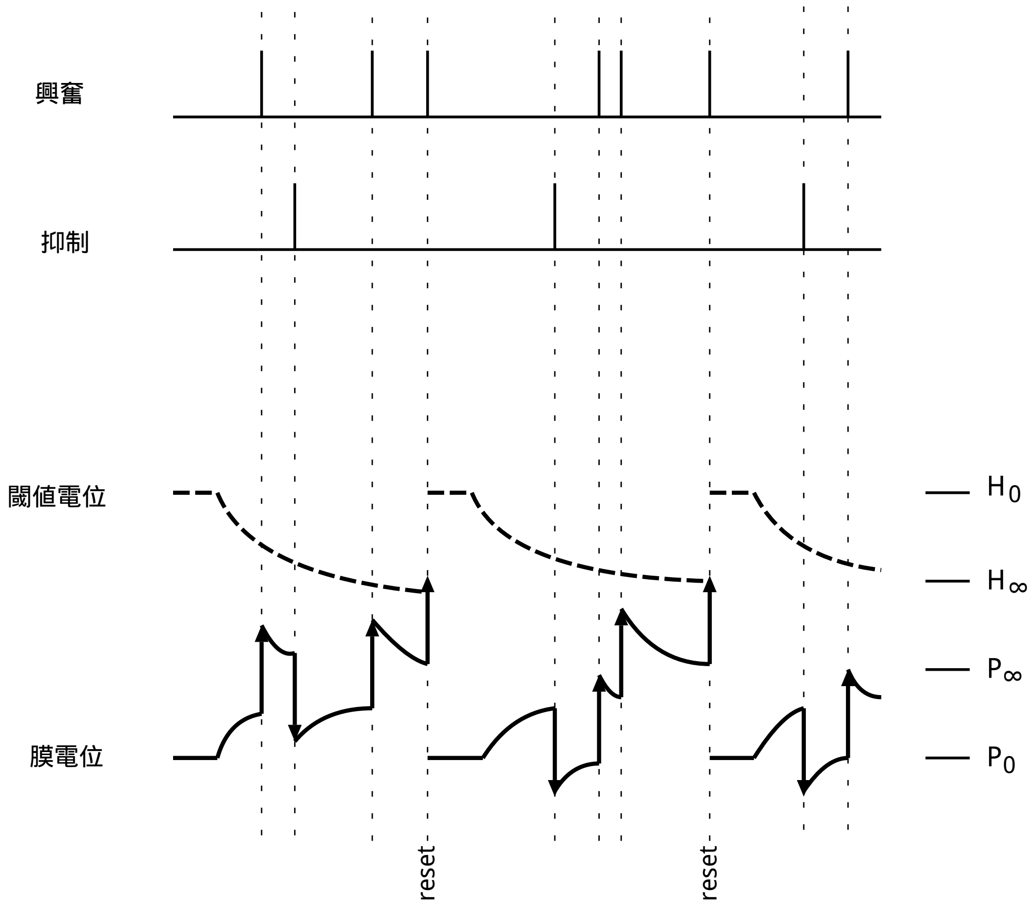

図 8: 入力の変化と膜電位の変化

平均頻度で表現するために実際の神経細胞における信号伝達に見られるパルス情報が失われているという欠点があった。

このためパルス動作型のニューロンモデルが用いられるようになってきている。このモデルはパルス到着の時刻や時間間隔の情報をも考慮した作りとなっているため、従来のモデルよりも高い表現能力を期待できる。しかし、このようなニューロンモデルによるネットワークの振舞いを調べるためには、従来のモデルとは異なり解析的に評価を行うことが困難である。よって、シミュレーションによってこの振舞いを調べる必要がある。

しかしこのモデルは構造や処理が複雑であるため計算量が膨大であり、高速のシミュレーションや神経回路網の大規模化は困難であるとされてきている。

このような高速シミュレーションの必要性から、最近では専用のプロセッサを作成することで 対処しようとする研究も行われている [11, 12]。ハードウェアでモデルを実現する場合は、ノードにおける計算、学習のための計算、接続

情報などを embedded にするかどうかが問題となる。

神経網をハードウェアで実現 (VLSI 化) することは、処理速度の向上を目的としているが、RISC の汎用ワークステーションが高速化している現在では、コストパフォーマンスの面では大幅に劣ると考えてよい。また速度の面においても、並列化および後述する Visual Instruction Set などを利用すれば、ソフトウェアシミュレーションもハードウェアでの実現速度に殆ど劣らないと考えられる。

しかし本質的な問題はその他にある。シミュレーションを行うにあたって最も重要なことは、まだ神経網のモデルが確定していないということにある。前にも述べたように、シミュレーション結果をモデルの設計へとフィードバックする必要がある状態である現在、シミュレーションをハードウェアで実現することは、モデルの可塑性を低め、最終的に adhoc なモデルへと行きついてしまう可能性が高い。また、ネットワークの構成が動的に変化するようなモデルを実現したい場合も、対処が困難であろう。

以上のような理由により、今回ソフトウェアによるシミュレータを実装し、その評価を行った。

5.3 シミュレーション方式

今回のニューロンモデルは、入力パルスを受けて動作するというパルス動作型ニューロンモデルと基本的に同じ構造をしている。

本モデルでは、一度 神経細胞が発火するとその膜電位 (興奮状態) はすぐにおさまるのではなく、徐々に減衰していくと考える (図 8)。このようなモデルで、毎回その膜電位を計算していたのでは、非常に計算コストが大きくなり効率が悪い。

また、細胞終端のシナプスに信号が到達するまでの遅延時間 (細胞間距離を表すもの) も各細胞ごとに設定されているため、これも考慮しなくてはならない。

しかし、そもそも各細胞における計算の目的は、その膜電位が閾値を超えるかどうかを確認することである。そこで、ある細胞の膜電位が閾値を超えるのがどういう場合かを考えてみると、膜電位は入力パルスの荷重和で与えられるので、入力パルスが変化したときであることが分かる。神経細胞の膜電位は単調減少しかしないので、ある神経細胞の入力へ繋がる神経細胞の少なくともひ

とつが発火したときには膜電位の計算を行う必要があるが、入力へ繋がる神経細胞の全てが発火していないときは、入力の全ては前時刻よりも減少しているはずであるから膜電位の計算を行うまでもなくその神経細胞は現時刻では発火しないことが分かるのである。

また、この方法では変化した信号が明らかであるため、減衰の式とパルス到着時刻さえ得られれば、簡単に現時点での入力値を計算できる。

よって、ある細胞が発火したときにその出力が繋がる先の細胞を計算のためにスケジューリングするという動作を繰り返せば、必要な神経細胞の計算は全て行えることになるのである。

以下に、具体的な離散事象シミュレーションの方法を述べる。ある神経細胞に対する信号の到着および膜電位の変化をひとつの事象として扱うこととする。神経細胞への信号の到達をイベントとするイベント駆動方式である。

まず考えられる方法としては、Time Wheel を用いることが挙げられる。ある神経細胞が発火すると、その出力が繋っている全ての神経細胞に対して事象を発生させる。事象は Time Wheel に登録されるため、全ての神経細胞に対する事象が一括管理できる。また、Time Wheel への登録の際に、各軸索の遅延時間を考慮し、それぞれに対応した異なる時刻エントリに事象の登録を行うことで、遅延時間の問題も簡単に処理できる。

Time Wheel は、軸索の最大遅延時間を超える長さであればよく、リング状に構成することによって、限られたメモリ空間においても無限時間のシミュレーションが可能である。

各神経細胞は、

- 当該細胞に信号が到着した時刻のうち、最も新しいもの
- そのときに変化した膜電位および閾値電位
- 膜電位、閾値電位の時定数
- 当該細胞の出力の結合先
- 結合それぞれに対する荷重および遅延時間

などを情報として持つ。直前の信号到着時刻と現時刻から経過時間を算出し、その経過時間を以前の膜電位の値からの減衰の計算に用いることで、現時刻の膜電位の変化の計算を行う。

図 9のようなネットワークにおいて、今 $t = 0$ で n_0 の出力が変化し、n_1 が発火したとする。このとき n_1 の発火により影響が及ぶノードは n_3, n_4, n_5

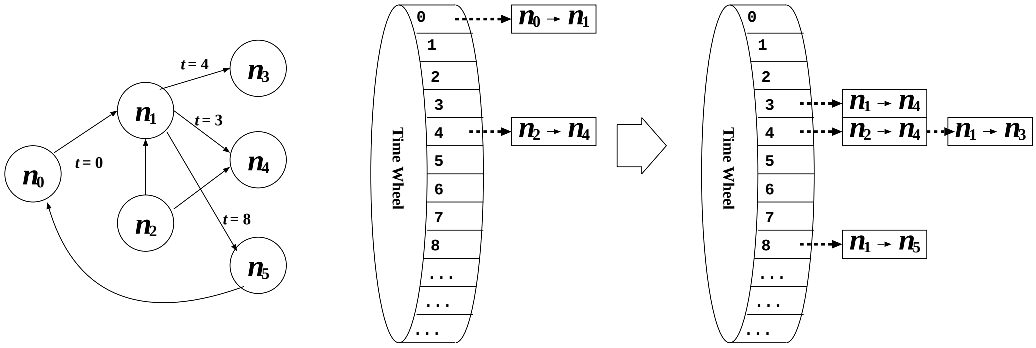

図 9: Time Wheel による事象管理

の 3 つであり、これらを Time Wheel に登録しなければならない。ここで n_1 と、その出力が繋がる先である n_3, n_4, n_5 との間の各エッジには、それぞれ遅延時間が設定されている。これは実際の神経網における、ノード間距離などをパラメータ化したものである。図 9の例では n_1, n_3 間のエッジの遅延時間は 4 となっているので、 n_1 の出力が n_3 に到達する事象を表すエントリを Time Wheel の $t = 4$ の箇所に登録する。 n_4, n_5 についても同様に、各エッジの遅延時間を考慮しながら Time Wheel への事象の登録を行う。

信号が到着したときに、各神経細胞で行うべき処理をまとめると以下のようになる。

1. 現時刻が、当神経細胞の不応期を経過しているか否かを判断する。神経細胞には一度発火するとしばらくは発火しないという不応期が存在し、これを考慮する必要がある。
2. 膜電位および閾値電位の計算
   (a) 不応期間中であった場合は、何も行わずに次の細胞に処理が移る。
   (b) 不応期が経過した場合は、当該神経細胞の膜電位および閾値電位を計算する。
3. 膜電位 と 閾値電位の比較 (発火の有無の判定)
   (a) 膜電位 $<$ 閾値電位 の場合、発火は起こらず次の処理に移る。
   (b) 膜電位 $>$ 閾値電位 の場合、
      i. まず、当該発火に伴い発生する不応期の、終了時刻を計算し登録しておく。現時刻から不応期時間後の時刻が、求める時刻である。
      ii. 当該神経細胞の出力が繋がる全ての神経細胞に対して、対応する軸索の遅延時間をテーブルから求める。タイムホイル上の、現時

刻からその遅延時間だけあとの時刻に対応するスロットに、事象を登録する。

ただし、この方法ではステップ時間の遅延しか実現できず、汎用性に欠ける。よって、実数時間の遅延も扱えるように事象管理の Time Wheel の構成を拡張した。

実数時間を扱う場合、スロットの数は固定数にできない。このため Time Wheel の環構造を 単なる Time Map に変更した。また、エントリの挿入 / 削除を効率よく行えるようにするため、 Time Map は binary tree 構造で実現した。

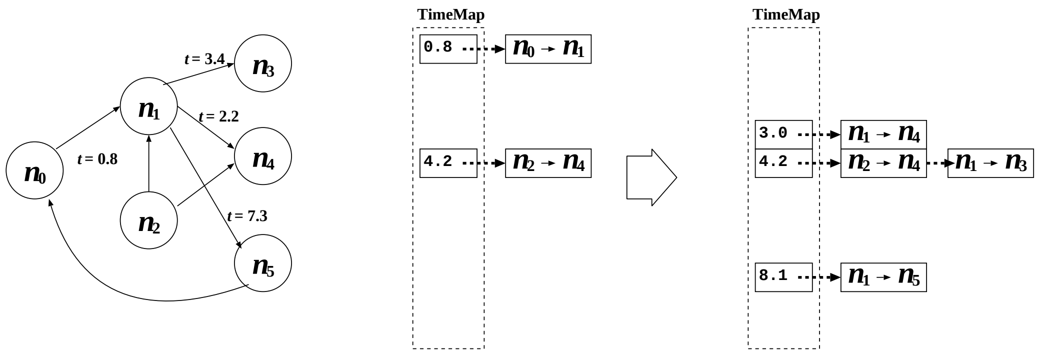

図 10: Time Map による事象管理

5.4 評価

本章で述べたパルス動作型ニューロンモデル用のシミュレータを実装し、評価を行った。

神経細胞数を 10000、各神経細胞あたりのシナプス結合数を 100 としてシミュレーションを行った。その結果を図 11に示す。横軸に神経細胞の発火率を示し、縦軸にはシミュレーションにおける単位時間の 10 ステップに対応するだけのシミュレーションに要した実時間を示した。ここでいう発火率とは、短い時間において神経網全体の神経細胞のうち発火している神経細胞が占める割合を指す。

図には比較のため、離散時刻事象シミュレーションを用いずに全ての計算を行った場合の平均所要時間も掲載した。この場合全てのステップにおいて全ての細胞の計算を行うため、所要時間が発火率に左右されず、ほぼ一定の値をと

る。ただしこの方法はステップ時間において各計算を行うため、軸索における遅延時間にはステップ時間の定数倍のみしか指定できない。

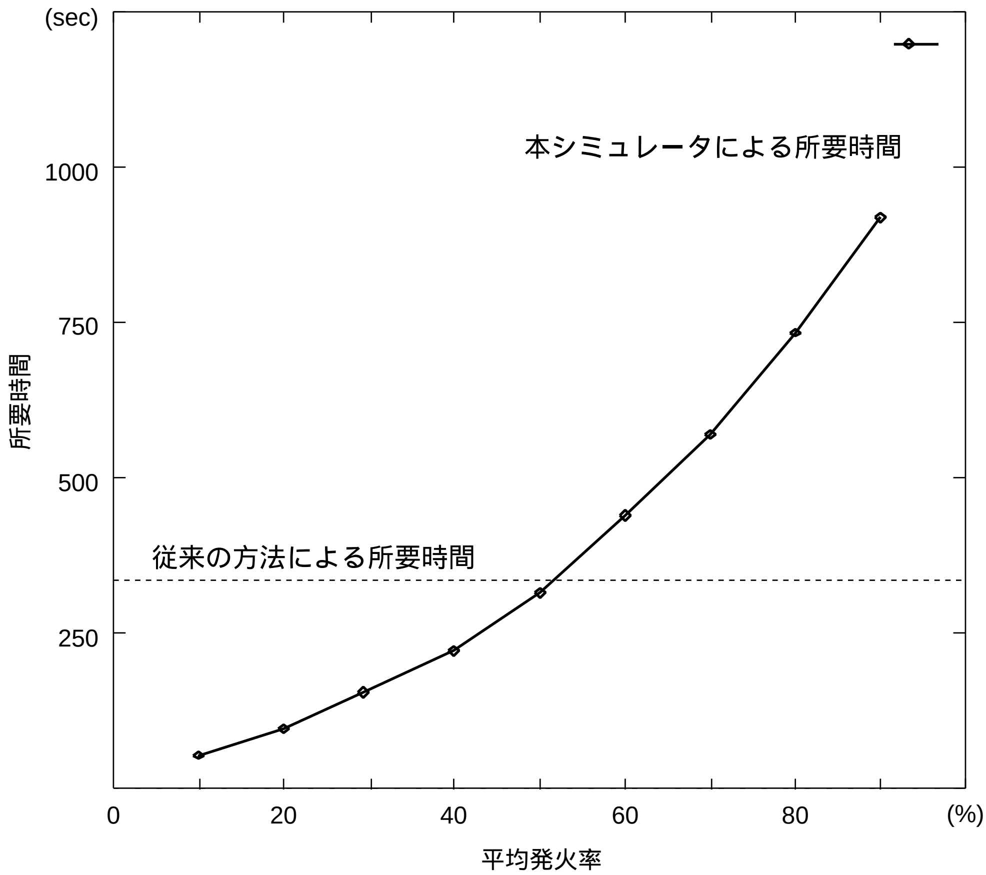

図 11: シミュレータの処理時間

本シミュレーションによる方法は、実数時間の遅延が扱える点において従来の方法よりも汎用性が高いという利点がある。にもかかわらず、図 11に示すように、発火率 10% の状態では約 4.4 倍 の速度向上が得られた。また、発火率が 50% を超えない範囲においては従来の方法よりも高速にシミュレーションが行えることが分かった。

人間の脳において、神経細胞の発火率は 10% くらいであるともそれ以下であるとも言われている。我々はひとつのモデルで脳全体のシミュレーションを行うわけではないため、脳の一部に注目した場合、発火率はもう少し高い値になると考えられる。しかし発火率 50% というのはかなり大きな値であり、通常のモデルでは発火率がそこまで高い値に及ぶことはほぼないであろうと思わ

れる。

つまり、従来の方法に較べてそのシミュレーションに要する時間を抑えることができるうえ、実数時間の遅延が扱えることでモデルの表現能力をより引きだせるシミュレーションが行える。

5.5 改良の余地

本シミュレータにおいて更なる高速化を図るため、シミュレーションの並列化方式について述べる。また、VIS の利用可能性についても触れる。

5.5.1 並列化方式

更に高速化を図るため、並列処理方式を提案する。Time Wheel を用いたイベント駆動方式を並列処理で行う場合、大きく分けてふたつの方法が考えられる。ひとつめの方法は、シミュレータ全体に対して Time Wheel をひとつだけ用意し、全ての事象の時刻同期を一括管理するという方法である。もうひとつの方法は、各プロセッサ (以下 PE と記述) にひとつづつ Time Wheel を持たせ、PE 毎に時刻同期をとりつつ処理を行わせる方法である。

前者の方法では、ひとつの Time Wheel によって全ての事象の管理を行うため同期を容易に実現できるという利点があるが、Time Wheel への事象の登録/参照のときに競合の生じる可能性があり、これを避ける手段を用意する必要がある。PE 台数が増加した場合、この競合のための排他制御のオーバヘッドと、排他制御により生じる待ち時間のため、十分な効果が得られない可能性もある。

後者の方法では、上記のような競合の起こる可能性は少ないが、各 PE 上の Time Wheel の間で動作の同期をとる必要がある。また、各 PE に Time Wheel を持たせる場合は、神経細胞を PE に静的に割当てることになるので負荷のばらつきが生じ、Time Wheel 間の同期によって待ち時間が増大してしまう可能性がある。

このようにふたつの方法にはそれぞれに長所 / 短所があるが、前者の競合回避のオーバヘッドがかなり大きくなると考え、以下では後者である各 PE に Time Wheel を持たせる方法で考える。ただし、並列化を行う場合 同期をとるのが困難であるため、遅延時間の設定に制限を与え、実数時間ではなくステップ時間による遅延を採用することとする。

具体的な並列化手法を以下に述べる。

全神経細胞を最初に各 PE に静的に分配しておき、各 PE の Time Wheel にはそれらの神経細胞に対応する事象のみの管理を行わせることとする。各 PE は、自分の Time Wheel から事象を取り出して処理し、その結果 新たな事象が発生すると、その事象に対応する神経細胞を担当する PE の Time Wheel に事象を登録する。

ここで、 PE が事象を参照するのは自分が管理する Time Wheel に対してだけであるから競合は生じないが、登録は他の Time Wheel に対しても行うためそのままでは競合が生じる。よって各 Time Wheel の各スロットには PE 数だけのフィールドを用意し、それぞれ PE は異なるフィールドを利用することでこの競合を避けることにする。

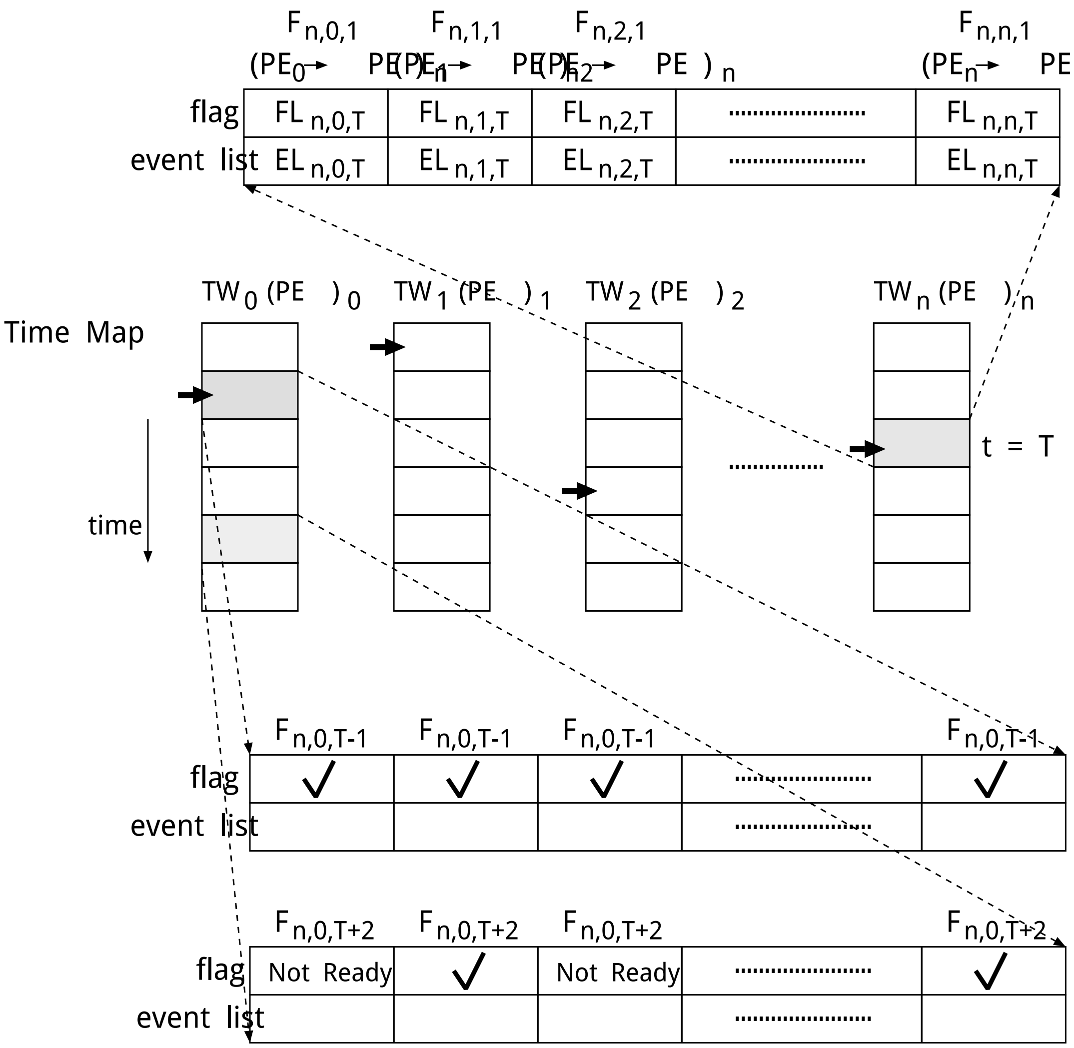

図 12: 並列化手法におけるタイムマップの構造

このときの各 Time Wheel の具体的なデータ構造は、図 12のようになる。PE 数と同じ数だけの Time Wheel、 TW_0 ... TW_n が存在し、それらは最初に各 PE に割当てられる。 PE_n の Time Wheel、 TW_n の時刻 t に対応するひとつのスロットには、やはり PE 数と同じ数の独立したフィールドが存在する。以下、 Time Wheel を TW_n、その時刻 t における PE_i からの信号に対応するフィールドを、 $F_{n,i,t}$ と表記することにする。

各フィールド $F_{n,i,t}$ はフラグ $FL_{n,i,t}$ と、事象リストへのポインタ $EL_{n,i,t}$ から成る。 $FL_{n,i,t}$ は、時刻 t までに PE_i に登録すべき事象を PE_j が全て登録しおわっているか否かを示すフラグであり、 $FL_{n,i,t}$ が立ったあとには それに対応する事象リスト $EL_{n,i,t}$ は変更を受けない。

プロセッサ PE_j において、時刻 t の処理が全て終了すると、$FL_{j,0,t}$ ... $FL_{j,n,t}$ をスキャンしてフラグが立っているものを調べ、フラグの立っているものに対しては それに対応する事象リストから全ての事象を順に処理する。事象が登録されているフィールド F に関しても、フラグ FL が立っていない場合は処理を行わない。

あるフィールドに存在する事象のリストを全て処理を終えると、フラグを off に変更し、次のフィールドのフラグを調べる。

$EL_{j,0,t}$ ... $EL_{j,n,t}$ にたいして全ての処理が完了すると、時刻 t を次に進めて同様の処理を繰り返す。またこの際、プロセッサ PE_j における Time Wheel 更新が完了したことを知らせるため、関係する Time Wheel のフラグフィールド FL を全てセットしておく。

このような方法をとることにより、各 Time Wheel において登録が完了しているか否かを PE が判断することができ、時刻の逆転が起こらずに並列に処理を実行できる。

5.5.2 Visual Instruction Set の利用

UltraSPARC の Visual Instruction Set [13] を利用することにより、シミュレーションの高速化を図る。

Visual Instruction Set(VIS™) は、汎用プロセッサ Sun UltraSPARC™ においてグラフィックス処理のための計算などを高速に実行するために用意された RISC 命令セットであるが、一部の基本演算を並列に行えるため数値解析や流体力学計算などの膨大な計算量を必要とするアプリケーションにも効果的に利用することができる。

UltraSPARC の中心部は 2 つの独立な ALU および ひとつの FGU (floating-point/graphics unit) から構成されており、更に FGU は、以下の 5 つの異なるユニットから構成されている (図 13)。

- floating-point divide/square root
- floating-point addintion/subtraction/absolute value/negative
- floating-point multiplication
- graphics addition, align, merge, expand, and logical
- graphics multiply, pack, compare and pdist

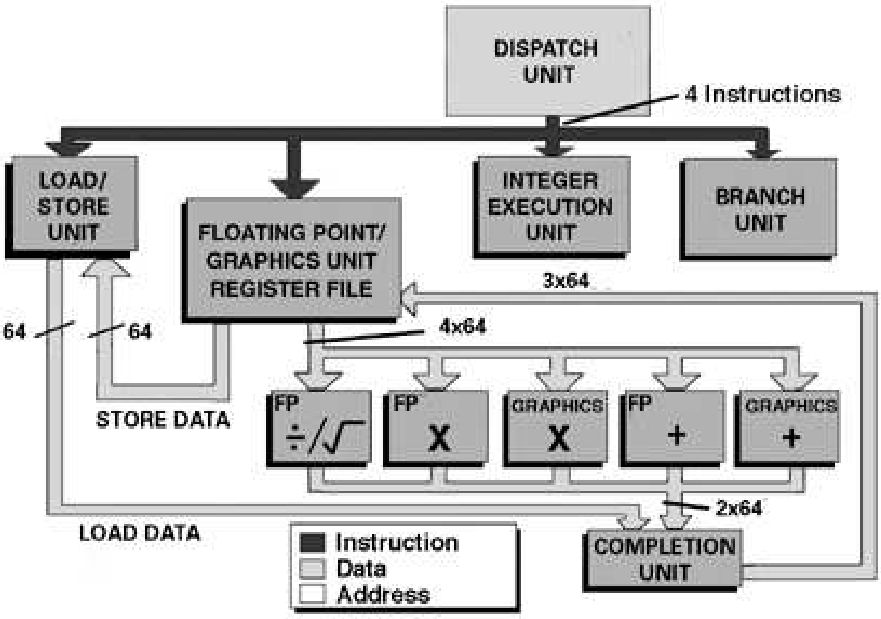

図 13: UltraSPARC の FGU と周辺

VIS ではグラフィックス処理に用いられる基本演算の高速化が図られている。例えば 16bit の値と 8bit の値の乗算では、1 サイクルで 4 つ の乗算が同時に行える。

神経網のシミュレーションでは、各神経細胞における入力の減衰計算は浮動小数点の乗算などを多く含み、最も処理に時間を要する部分である。よって、これらの演算命令をこの減衰計算部に用いることで、更なる高速化が期待できるであろうと考える。

6　おわりに

本論では、これから行う人間のように知的なシステムを構築することを目標とする研究の準備として、その際に必要となると考えられる神経網のシミュレーションを高速に行えるためのシミュレータを開発し、その評価を行った。

我々の研究の方向性を明確にするため、人間のように知的なシステムを構築するという目標に至るための別のアプローチとして 人工知能研究をとりあげ、その表象性、明示性、目的性において 我々のアプローチとの比較を行った。

人工知能研究が、知識や処理の記号化という高い表象性を前提としたアプローチであるのに対し、我々の研究がその記号化自体を研究の対象とすることを示した。またそのために、最近研究がすすんでいる人間の脳に対し、記号化の方法の手本を求めることにした。

脳の情報処理において鍵となるのは、その表象化の方法すなわち記憶のメカニズムであると考え、現在知られている脳の記憶の方法について概観した。

以上のことを踏まえたうえで、脳の記憶構造を探るために神経網のシミュレーションを行う必要性について述べた。また、モデルが決定しない現在において、ハードウェアではなくソフトウェアでシミュレーションを行う必要性について述べた。

このソフトウェアシミュレーションに用いるための高速シミュレータを開発し、その評価を行った。計算量の低減には、LSI の論理シミュレーションなどで用いられる離散時刻事象シミュレーションの手法を応用し、単位時間に限らず実数時間の軸索遅延が扱えるように拡張を行った。

このシミュレータを用いて評価を行った結果、従来の方法に較べて実数時間の遅延が扱える点において汎用性が高いにもかかわらず、神経細胞の発火率が 10% の状況においては従来の方法の 約 4.4 倍 の速度向上が得られることを示した。また、神経細胞の発火率が 50% を超えない範囲においては従来の方法よりも高速にシミュレーションが行えることを示した。

また、シミュレータの更なる高速化のために、その並列化手法について考察を行った。各 PE が持つ Time Wheel の各時刻スロットに、PE の総数だけのフィールドを用意することで、競合を避けながら PE 間の同期をとることができることを示した。

謝辞

寛大なお心をもって本研究の機会を与えて下さり、適切な御指導を賜わりました富田眞治教授に深甚な謝意を表します。

また、貴重な御助言をいただいた森眞一郎助教授、五島正裕助手に深く感謝致します。

さらに、日頃暖かく御鞭撻下さった京都大学大学院工学研究科情報工学専攻富田研究室の諸兄に感謝致します。

最後に、今まで私に何らかの形で関わりのあったすべての方々に感謝致します。ありがとうございました。

参考文献

[1] 梶井基次郎: 檸檬.

[2] Fodor, J.: The modularity of mind., *MIT Press* (1983).

[3] Turing, A. M.: Computing machinery and intelligence, *Mind*, Vol. LIX, No. 236 (1950).

[4] Ogawa, S., Tank, D. W., Monon, R., Ellermann, J. M., Kim, S. G., Merkle, H. and Ugurbil, K.: Intrinsic Signal Change Accompanying Sensory Stimulation: Functional Brain Mapping with Magnetic Resonance Imaging, Vol. 89, USA, Proc. Natl. Acad. Sci., pp. 5951–5955 (1992).

[5] Tovée, M. J. and Cohen, E. M.: Working memory: Trouble in mind, *Current Biology*, Vol. 6, No. 13 (1996).

[6] Bliss, T. V. P. and Lϕmϕ, T.: Long-lasting potentiation of synaptic transmission in the dentate area of the anesthetized rabbit following stimulation of the perfoant path, *Journal of physiology*, Vol. 232, No. 31 (1973).

[7] 沢田康次: 脳の理工学的研究の方法論について, 日本神経回路学会誌, Vol. 3, No. 4, 日本神経回路学会, pp. 123–125 (1996).

[8] : 脳型コンピュータとは何か, 日本神経回路学会第 8 回全国大会予稿集 (1997).

[9] Hebb, D. O.: The problem of consciousness and introspection, *Brain*

mechanisms and consciousness (Delafresnaye, J. F.(ed.)), Blackwell, Oxford (1954).

[10] Segundo, J. P., Parkel, D. H., Wyman, H., Gegstad, H. and Moore, G. P.: Input-output relations in computer-simulated newve cells, *Kybernetic*, Vol. 4, No. 5, pp. 157–171 (1968).

[11] 市川道教: Introduction of BMP chip design 脳模倣プロセッサーの概要, *Brain Function & Brain Computing Workshop*, Vol. 161, Tsukuba City, Japan, FED, pp. 48–49 (1998).

[12] 松本元, 重松征史, 市川道教: 脳研究から情報処理へ – 脳型コンピュータの開発にむけて –, 情報処理学会誌, Vol. 38, No. 5 (1997).

[13] Sun Microsystems, Inc.: *ULTRASPARC™ THE VISUAL INSTRUCTION SET (VIS™): ON CHIP SUPPORT FOR NEW-MEDIA PROCESSING*, 901 San Antonio Road, Palo Alto, CA 94303 USA (1995).