

Title	<特集：モデル> フォン・ノイマンとマカロック ピッツ ・モデル：オートマトン理論の誕生
Author(s)	伊藤, 和行
Citation	科学哲学科学史研究 (2008), 2: 117-132
Issue Date	2008-01-31
URL	https://doi.org/10.14989/56984
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

フォン・ノイマンとマカロック - ピッツ・モデル

オートマトン理論の誕生

伊藤和行*

Abstract

Von Neumann's "First Draft of a Report on the EDVAC" is known as the first paper on the modern computation theory, where he compared electric circuits to neuron networks. The idea of conceiving neuron networks as logical circuits was proposed by W. S. McCulloch and W. Pitts, and was called the McCulloch-Pitts Model. This model led von Neumann to develop his automata theory, in which he proposed five models like the cellular model. Through them, he tried to present a unified study of modern computing machines and the human nervous system.

1. 序

デジタル・コンピュータの起源をめぐることは、これまで様々な議論がなされてきたが、一般には、第二次世界大戦中にペンシルバニア大学ムーア・スクールで開発された ENIAC と言われている。しかしこの見解にもいくつかの異論が出されている。まず ENIAC の開発の中心の一人だったモークリーが、その理論的アイデアを、アイオワ州立大学のアタナソフが開発していた ABC マシンから得ていたということがある^{*1}。また ENIAC 自体、我々の抱くデジタル・コンピュータとはいくつかの重要な点で異なっていた。第一に、プログラムを電気回路によって物理的に実現しており、プログラムをメモリ上に保存する「プログラム内蔵」コンピュータの起源とみなすことはできないという点が挙げられる。またさまざまな計算を迅速に行なうために並列処理を行なうとともに、それまでの計算機の慣習に倣い 10 進法による数値計算も行なっている。いかなるマシンが最初のデジタル・コンピュータなのかという問題は、何をもってデジタル・コンピュータとみなすのか、デジタル・コンピュータの本質をどこに見いだすのかという問題であると言うことができる。前者の問いの解答は、後者の問いの解答次第と言えるだろう^{*2}。

以上のように、デジタル・コンピュータの起源の問題は厄介な側面を抱えているが、それに対してコンピュータ理論の起源に関しては、そのような問題はないのではないだろうか。第一に挙げられるのは、いわゆる「ノイマン型コンピュータ」の基本的な発想を展開したフォン・ノイマ

* 京都大学大学院文学研究科

^{*1} バークス『誰がコンピュータを発明したか』、大座畑重光監訳、マッカーズ共訳、工業調査会、1998 参照。

^{*2} この起源の問題に関する明快な議論が、星野力『誰がどうやってコンピュータを創ったのか?』(共立出版、1995)に見いだされる。

ン (John Von Neumann, 1903-1957) の論文「EDVAC に関する報告第一草稿」(“First Draft of a Report on the EDVAC”) である^{*3}。この論文は、命令をデータとともにメモリに入れ、その命令を逐次的に処理するという「可変プログラム内蔵」方式が提唱されていることで知られ、コンピュータの歴史書を開くと必ず言及されるものである^{*4}。この論文に実際に眼を通すと、プログラム内蔵に関する記述は意外と短く、メモリに係わる記述が後半部分の大半を占めていることが目に付く。これには、当時メモリの開発が大きな問題となっていことが背景にある。またフォン・ノイマンのコンピュータへの関心が、第一に流体力学を中心とした数値計算にあり、その計算の途中結果をどのようにメモリ上に置くのかということが計算精度と密接に結び付いていたことがあった。

この論文の特徴は、個々の技術的な細部を抽象し、論理的構造を明らかにした点であると言われる。電子素子から構成される論理回路による数値計算の実現はこの論文の主要なテーマの一つであるが、そこでは、コンピュータの構成する電子素子とニューロン（神経細胞）との類比という我々には馴染みのない話題にかなりの紙面が割かれていた^{*5}。デジタル計算機の計算装置は「素子」(“element”) から構成されている。「素子」とは、二つあるいはそれ以上の不連続な平衡状態を持つ「リレー」(“relay”) であり、それ自身でその状態を保持できるが、外的な刺激によってある状態から他の状態に移ることができ、それに対応した刺激を外部に出すことができる。「素子」としては円盤のような機械的なもの、電信リレーのような電気機械的なもの、そして真空管のような電子的なものが用いられていた。フォン・ノイマンによれば、ニューロンは、上で定義された意味において「素子」なのである。というのは、ニューロンは「沈黙」と「興奮」という二つの異なる状態を持ち、他のニューロンと結合していて、興奮したときには刺激を発するからである。現実のニューロンは、閾値、時間的総量、相対的抑制などの様々な機能を持っているが、それらが捨象されて単純化されたものの機能は真空管によって模倣可能であると彼は考えていた。

この計算機と神経系における「素子」の類比というテーマは、彼のオートマトン^{*6}の理論の出发点となったことが知られている。特定の入力に対して一定の出力を示す自動機械、さらには自己複製機械としてのオートマトンの概念は、機械と人間を情報処理という統一的な視点から捉える試みであり、サイバネティクスの理論などとともに現代の情報理論の起源の一つであったと言

^{*3} もちろん計算機械の理論的な基礎付けに関するチューリングの論文「計算可能な数と決定問題への応用」(“On Computable Numbers, with an Application to the Entscheidungsproblem”, 1936) を忘れることはできないが、ここでは現実的なコンピュータの理論ということでフォン・ノイマンの論文のみを挙げている。チューリングに関しては、星野力『甦るチューリング：コンピュータ科学に残された夢』、NTT出版、2002などを参照。

^{*4} 「草稿」は以下の版で読むことができる。W. Aspray and A. Burks, eds., *Papers of John von Neumann on Computing and Computer Theory*, Cambridge, Mass, 1987, pp.17-82; “First Draft of a Report on the EDVAC”, ed. by M. D. Godfrey, *IEEE Annals of the History of Computing*, 17 (1993), pp.7-75. その内容に関しては、小山俊士「フォン・ノイマン「EDVAC 草稿」について」、『科学史・科学哲学』, No.19, 2005, pp.106-123 を参照。

^{*5} 「草稿」は15の章から構成されているが、「素子」について論じているのは、第4章「素子、同期、ニューロンとの類似」と第6章「E素子」である。

^{*6} 単数形では“automaton”, 複数形では“automata”となるが、本稿では単数形の「オートマトン」に統一する。

われる。

この神経回路を論理回路として捉える試みは、フォン・ノイマンのオリジナルではなく、マカロックとピッツが1936年に発表した論文「神経活動に内在する観念の論理計算」で展開したものである。この論文はニューロン・ネットワークに関する最初の論理的考察として知られ、そこで提出された神経回路に関するモデルはマカロック・ピッツ・モデルと呼ばれている。以下ではこのモデルを紹介するとともに、それから得られた着想をフォン・ノイマンがオートマトンの理論へとどのように発展させていったのかを辿ることにしよう。

2. マカロック・ピッツ・モデル

マカロックとピッツの共著論文「神経活動に内在する観念の論理計算」(“A Logical Calculus of the Ideas Immanent in Nervous Activity”, 1943)^{*7}は、神経回路網に関する最初の数学的モデルを提示したことで知られる。そこでは分析を容易にするために、ニューロンの入力と出力を二値に限定し、また時間的遅延を離散的に扱うという制限がなされ、きわめて単純化されたモデルであったが、神経回路網にとってもっとも本質的なことはこのモデルによって表現されており、以降の研究の出発点となったと言われる^{*8}。

著者であるマカロックとピッツは、非常に異なる学問的バックグラウンドを持っていた。マカロックは(Warren Sturgis McCulloch, 1898-1969)神経生理学者として活動し、彼の研究目標は「実験的認識論」、すなわち脳は、知覚のような心的な現象を維持するためにどのように機能するのかを実験的に解明し、心の機能を脳の中の神経系の過程に統合することであった^{*9}。彼は、1930年代にはチンパンジーとサルを飼った実験を行ない、大脳皮質において電気パルスの伝播経路について研究していた。さらに彼は、認識過程のメカニズムを明らかにするために、大脳皮質内の神経の機能構造の実験的研究に、形式論理学を組み合わせることを考えたのだった。しかしマカロックは数理論理学には疎く、このプログラムを進めるためにはその分野に秀でた協力者が必要だった。それがピッツ(Walter Pitts, 1923-1969)である。彼は神童と呼ばれ、15歳のときにカルナップから数理論理学を、またラシェフスキーから数理生物学を学んでおり、17歳のときにマカロックに会ってからは彼の共同研究者となった。ピッツの数理論理学者としての優れた才能は、マカロックの神経系に関するプログラムの突破口となった。また彼はウィナーの共同研究者としても活動し、サイバネティクスの誕生に際して重要な役割を果たしたと言われる。

マカロックとピッツの共著論文「神経活動に内在する観念の論理計算」は生物系の要素に数理

^{*7} McCulloch, W. S. and W. Pitts. “A Logical Calculus of the Ideas Immanent in Nervous Activity”, *Bulletin of Mathematical Biophysics*, 5 (1943), pp.115-133; McCulloch, *Embodiments of Mind*, Cambridge, Mass., 1965, pp.19-39.

^{*8} ガードナー『認知革命 - 知の科学の誕生と展開』, 海保博之・佐伯胖監訳, 産業図書, 1987, pp.17-18.

^{*9} マカロックに関しては、ハイムズ『サイバネティクス学者たち アメリカ戦後科学の出発』, 忠平美幸訳, 朝日出版社, 2001, 第三章, 49-77頁; Abraham, T. H., “Integrating Mind and Brain: Warren S. McCulloch, Cerebral Localization, and Experimental Epistemology”, *Endeavour*, 27-1 (2003), pp.32-36.

論理学を適用した最初の試みである^{*10}。それは、個々のニューロンを、複数の入力に対して定まった論理規則に従って出力を与える論理素子とみなし、神経回路網を論理回路として捉える試みだった。彼らは神経刺激に対して0と1という二値のみを考え、興奮ニューロンと抑制ニューロンからなる様々な結合を持つ仮想的なネットワークを構想し、それと命題論理との同型性を主張している。彼らの行なった単純化のために、彼らのモデルは、当時の神経生理学者にはほとんど反響をもたらさなく、むしろフォン・ノイマンが「草稿」において取り上げられたことによって世の中に知られるようになり、サイバネティクス運動の中で重要な役割を果たすことになったのだった。

彼らは、まず序論において、神経回路とニューロンに関して当時一般的に認められていたことをまとめている。神経系はニューロンの回路網であり、各ニューロンは細胞体と軸索からなっていて、あるニューロンの細胞体と他のニューロンの軸索との接触がシナプスである。ニューロンは発火と休止という二つの状態を持ち、前者の状態では他のニューロンへ刺激を発する。発火の状態になるのは、他のニューロンからの刺激が閾値を超えたときである。各ニューロンは複数のニューロンとシナプスを通じて繋がっており、その状態は他のニューロンからの刺激によって決定される。シナプスには、興奮を引き起こすものと興奮を抑制するものの二種類がある。またある時間間隔における刺激の時間的な総和が閾値を超えると発火する場合もある。

以上のようなことに加えて、著者たちは5つの「計算にもっとも便利であるような物理的な仮定」を置いていた。それらは「複雑な回路の振舞を命題記号論理の表記で記録」^{*11}するための仮定である。

1. ニューロンの活動は「全てか無か」の過程である。
2. 任意の時刻に一つのニューロンを興奮させるためには、潜在的な加算時間においてある定まった数のシナプスが興奮しなければならない。
3. 神経系内の唯一意味のある遅延はシナプス遅延である。
4. 任意の抑制シナプスの活動はその時刻におけるニューロンの興奮を絶対的に妨げる。
5. 回路の構造は時間とともに変化しない^{*12}。

しかしこの仮定は、いくつかの問題点を含んでいた。いうまでもなく、各ニューロンは異なる閾値を持っており、そしてニューロンの活動は「全てか無か」の過程でもなかった。また著者自身が指摘しているように、実際の回路は時間的な変化を伴っていた。第一に、促進や消去のように、時間的に先行する活動が、続く刺激に対する反応を変える現象があった。第二に、学習のよう

^{*10} この論文の背景と影響に関しては、Abraham, T. H. "(Physio)logical Circuits: The Intellectual Origins of the McCulloch-Pitts Neural Networks", *Journal of the History of the Behavioral Sciences*, 38 (2002), pp.3-25; Kay, L. E., "From Logical Neurons to Poetic Embodiments of Mind: Warren S. McCulloch's Project in Neuroscience", *Science in Context*, 14 (2001), 591-614.

^{*11} McCulloch and Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", pp.116-117.

^{*12} Ibid., p.118.

に、ある先行する時刻になされた活動が回路に影響に与えて変えてしまい、以前は適切な刺激が今では適切なものになるのである。そのような問題点に対して、著者は、ここでの形式的な議論は実際の説明とは考えていないことを強調している。ここでは「神経系の形式的な扱い」を目指すのであって、扱われるのは「命題論理」によって記述できる限りのものなのである。また著者たちは触れていないが、時間も離散的に扱うことが仮定されている。すなわち、ある時刻 t に先立つ時刻は、 $t-1, t-2, t-3$ 等となる。

実際の考察は「循環を含まない回路網」と「循環を含む回路網」に分けられ、各々5つの定理が置かれている。とくに後者はニューロンが円環上に繋がった経路を持つ回路であって、その活動は再帰関数によって表わされる。「循環を含まない回路網」の考察は、さきほど触れた5つの「物理的前提」から始まっているが、そこでは現在では用いられていない論理記号が使われ、そのことが理解を困難にしている^{*13}。その詳細の理解には、後にクリーネが現代的な記号を用いて分析した論文での議論を辿るのがよいように思われる。またクリーネは「循環を含む回路」の記述は曖昧であることを指摘しており、論文では、マカロックらとは異なる仕方で問題を扱っている^{*14}。

ここでは、フォン・ノイマンの論文でも似たような図がみられる、選言や連言、否定に対応するいくつかのニューロン回路の図を見てみよう^{*15}。ここで $N_i(t)$ はニューロン c_i の時刻 t における活動を示している。また以下の説明では、表記を現代のものに変更している。

$$a: N_2(t) = N_1(t-1)$$

c_2 が時刻 t において発火するのは、 c_1 が時刻 $t-1$ において発火するときのみである。

$$b: N_3(t) = N_1(t-1) \vee N_2(t-1) [\text{選言に対応}]$$

c_3 が時刻 t において発火するのは、時刻 $t-1$ において c_1 か c_2 が発火するときである。

$$c: N_3(t) = N_1(t-1) \wedge N_2(t-1) [\text{連言に対応}]$$

c_3 が時刻 t において発火するのは、時刻 $t-1$ において c_1 と c_2 の両方が発火するときである。

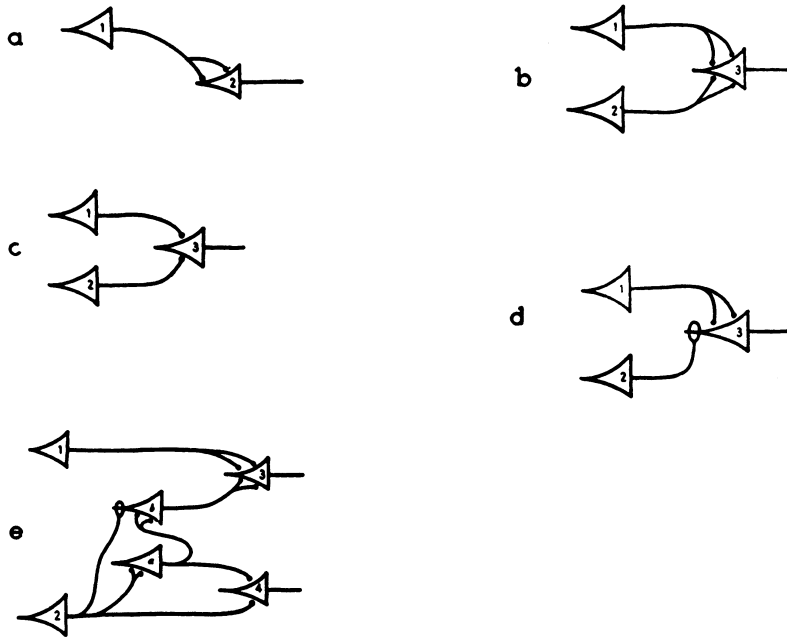
$$d: N_3(t) = N_1(t-1) \wedge \sim N_2(t-1)$$

c_3 が時刻 t において発火するのは、時刻 $t-1$ において c_1 が発火し、 c_2 が発火しないときである。

^{*13} そこで用いられていた表記法は、Carnap, R., *The Logical Syntax of Language*, 1938 と Russell, B., and Whitehead, A. N., *Principia Mathematica*, Cambridge, 1925 のものである。

^{*14} Kleene, S. C., "Representation of Events in Nerve Nets and Finite Automata", in C. E. Shannon and J. McCarthy, *Automata Studies*, Princeton, NJ, 1956, pp.3-41, spec. p.4.

^{*15} 論文には a から i まで 10 個の図があるが、すべての図が本文のときちゃんと対応しているわけではない。McCulloch and Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", p.130.



e 上: $N_3(t) = N_1(t-1) \vee [N_2(t-3) \wedge \sim N_2(t-2)]$

c_3 が時刻 t において発火するのは, c_1 が時刻 $t-1$ において発火するか, c_2 が時刻 $t-3$ において発火し $t-2$ においては発火しないときである.

e 下: $N_4(t) = N_2(t-2) \wedge N_2(t-1)$

c_4 が時刻 t において発火するのは, c_2 が時刻 $t-2$ と $t-1$ の両方において発火するときである.

論文本文の内容は完全に数理論理学のものであって、当時の神経回路の研究者はほとんど理解できなかったろうし、この神経回路を数理論理学で表現しようという試みの意図がわからなかったと思われる。著者たちは、神経生理学者からの批判を予期して、神経回路網の振舞と命題論理との「形式的」な同定性をここでは前提しているが、それを現実のものにまで拡張しようというものではないと述べていた^{*16}。数学を生命現象の分析に積極的に利用しようという試みは彼らが最初だったわけではなく、この論文が掲載された雑誌 *Bulletin of Mathematical Biophysics* の創刊者であるラシェフスキ (Nicolas Rashevsky, 1899-1972) がすでに神経生理学において試んでいた。しかしラシェフスキと彼らとは、そのアプローチに大きな違いがあった。ラシェフスキはニューロンの興奮と抑制を連続的な過程として捉え、それを分析するモデ

^{*16} Ibid., p.117. 彼らは、1947 年の論文 (“How We Know Universals: The Perception of Auditory and Visual Forms ”) において知覚作用における神経回路網の機能を論じた際には、この論文では捨象された性質を持つ、もっと現実的なものとしてニューロンを論じている。Cf. Abraham, T. H., “From Theory to Data: Representing Neurons in the 1940s ”, *Biology and Philosophy*, 18 (2003), pp.415-426.

ルでは微分方程式による解析的手法を用いていたのである^{*17}。マカロックとピッツが初めてニューロンの活動を「全てが無か」の過程であると仮定して、神経回路網のモデルに数理論理学を導入したのだった。

マカロック ピッツ・モデルは、現在では、神経科学においても神経回路網に関する先駆的な研究として評価されているが、それに続く研究が現われるのはかなり後のことであり、当時は研究者から無視されてしまった。世の中に知られるようになったのは、もっぱらフォン・ノイマンが講演や論文で紹介したことによるものであり、その影響が現われたのはオートマトン理論、サイバネティクス、人工知能という研究分野においてだった。またマカロック自身、ウィーナーやフォン・ノイマンと並んで初期のサイバネティクス研究において中心的な役割を果たしており、米国サイバネティクス学会の初代会長も務めている。彼らのアプローチは、生物と人工物を情報処理という共通の視点から捉えていこうと試みの先駆と見なされたのである。

3. オートマトン理論へ向けて

フォン・ノイマンが「草稿」において述べていたニューロンと電子素子との類比は、生物と人工物に関する情報処理システムという統一的な視点からの研究、オートマトン理論として展開されることになる。オートマトンに関する考察は「草稿」執筆とほぼ同じ時期に始まっているが、現在入手できる最初の論考が発表されたのは、1948年9月に行なわれた「ヒクソン・シンポジウム」においてである^{*18}。このシンポジウムは脳のメカニズムに関して、少数の神経生理学者や心理学者を集めて行なわれ、そこで彼は「アナログ網とオートマトンの論理」(“The Logic of Analogue Nets and Automata”)と題する講演を行なった。その内容は後に出版されたシンポジウムの報告書における論文「オートマトンの一般的かつ論理的理論」(“The General and Logic Theory of Automata”)から知ることができる^{*19}。そこでは、数理論理学に基づいた自然オートマトンと人工オートマトンについての理論の必要性、チューリングの研究とマカロック ピッツの研究の重要性が述べられており、研究すべき課題としてオートマトンに関する確率的理論、オートマトンの複雑性と自己複製、さらにはコンピュータと脳の比較が挙げられ、彼のその後の研究の方向性が示されている。

フォン・ノイマンは1957年に亡くなるまでに、この報告を含めてオートマトンに関連して5

^{*17} Aizawa, K., “Some Neutral Network Theorizing before McCulloch: Nicolas Rashevsky’s Mathematical Biophysics”, in R. Moreno-Díaz and J. Mira, eds., *Brain Processes, Theories, and Models: An International Conference in Honor of W. S. McCulloch 25 Years After His Death*, Cambridge MA, 1996, pp.64-70.

^{*18} 彼はすでに1946年6月にプリンストン大学で講義において自己複製の問題を論じていた。フォン・ノイマンのこの分野の活動については、アスプレイ『ノイマンとコンピュータの起源』、杉山滋郎・吉田晴代訳、産業図書、1995、第8章を参照。

^{*19} L. A. Jeffress, ed., *Cerebral Mechanics in Behavior: the Hixon Symposium*, New York, 1951, pp.1-41; A. H. Taub, ed., *Collected Works*, General Editor, A. H. Taub, 6 Vols., New York, 1961-1963; Vol. 5, pp.288-328.

つの論考を残している。

1. 「オートマトンの一般的かつ論理的理論」(1951, 発表は1948)
2. 「複雑なオートマトンの理論と組織化」(1949)
3. 「確率的論理学, および信頼度の低い部品から信頼度の高い有機体を形成すること」(1952)
4. 「オートマトンの理論: 構成, 複製, 均質性」(1952-)
5. 『コンピュータと脳』(1958, 執筆は1955-1956)^{*20}

これらの論考の中で2と4は, 残された原稿をアーサー・バークスが整理して解説を加えてまとめ, 『自己増殖オートマトンの理論』(1966)^{*21}として出版している。それらはフォン・ノイマンがオートマトンの理論を本格的に展開したものであるが未整理の状態であって, 彼の理論は未完成のまま残されていた。以下では, 「オートマトンの一般的かつ論理的理論」にそって, フォン・ノイマンがマカロック・ピッツの研究を契機として, オートマトン理論をどのように発展させていったのかを見ることにしよう。

「オートマトンの一般的かつ論理的理論」は, 導入部分と以下の7つの章から構成されている^{*22}。

1. 予備的考察
2. 計算機のいくつかの関係する特徴に関する議論
3. 計算機と生物体との比較
4. オートマトンの将来の論理的理論
5. デジタル化の諸原理
6. 形式的ニューロン回路網
7. 複雑性の概念; 自己複製

この論考を通じて, オートマトンの明確な定義は示されていないが, フォン・ノイマンは生物と人工物を情報処理という共通の視点から考察しようと試みている。自然のオートマトンとして考えられているのが生物の神経回路網とりわけ人間の中枢神経系であるのに対して, 人工のオートマトンとして考えられているのは電子計算機すなわちコンピュータである。両者のシステムを構成している素子はいうまでもなく大きさも素材もまったく違うものである。そこで彼は両者で相違のある素子の生理学的な側面は捨象して, 素子を外部との情報機能のみが定義された

^{*20} 1. “The General and Logic Theory of Automata”, 2. “Theory and Organization of Complicated Automata”, 3. “Probabilistic Logics and the Synthesis of Reliable Organism from Unreliable”, 4. “The Theory of Automata: Construction, Reproduction, Homogeneity”, 5. *The Computer and the Brain* [『電子計算機と脳』, 飯島泰蔵・猪俣修二・熊田衛共訳, ラティス, 1964.]。

^{*21} *Theory of Self-Reproducing Automata*, edited and completed by Arthur W. Burks, Illinois, 1966. [『自己増殖オートマトンの理論』, 高橋秀俊監訳, 岩波書店, 1975.]

^{*22} 原文には章番号がないが, 参照しやすいように本稿では章番号を付けてある。

「ブラックボックス」とみなすことを提案する．そしてどのように素子が全体として組織されるのか、どのようにして全体の機能が素子の機能から構成されるのかを理解しようというのである．このように素子の振舞を機能的な特性のみによって規定するアプローチは「公理的手続き」(“The Axiomatic Procedure”)*²³と呼ばれている．この方法でなされた過去の研究の中で最も重要なものが、マカロック - ピッツの研究だった．「公理論的手続き」によって自然と人工のオートマトンを論じる際には、両者のシステムを構成する素子数の違いに注意するのを忘れてはならない．人間の中枢神経系のニューロン数は 10^{10} のオーダーであって、このように複雑なシステムはこれまで考察されたことがないのである．一方人工オートマトンの構成部分の数は 10^3 から 10^6 のオーダーであり、そのプロトタイプとも言える計算機の場合は 10^3 から 10^4 程度であろう．

第2章以降では、生物体と計算機を比較しつつ、前者により近い機能を持つ計算機を可能にするための条件を論じている．第2章「計算機のいくつかの適切な特徴に関する議論」は、計算機の基本原理の紹介に割かれていた．これは、この論文の元となった発表が行なわれたヒクソン・シンポジウムの出席者が主として脳生理学者であったこと、当時デジタル・コンピュータの概念がまだ一般化していたことによるためであろう．計算の精度と信頼性がアナログ計算機とデジタル計算機に関して評価され、後者の優位性が主張されている．アナログ計算機は当時科学計算において一般的に用いられており、数を物理量、たとえば電流、電圧、円盤の回転角度などで表わすという原理に基づいている．既存のアナログ機械として最も優れていたのは、ブッシュ (Vannevar Bush, 1890-1974) が開発した「微分解析機」(Differential Analyzer) であって、様々な機械的機構を用いて、積分などの数値演算を実行し、グラフによって結果を表示した．アナログ計算機に対してデジタル計算機が優れている点として第一に挙げられるのは、「信号と雑音比」の問題である．すなわち機械が操作する数を表わす「信号」に比べて、「雑音」を構成する機構の制御不能な揺動をどれだけ低く抑えられるかという問題である．アナログ機械の「最大信号レベル」は、機械的機構では $1:10^4$ あるいは 10^5 であって、電気的機構では $1:10^2$ である．これらの比は、計算の初期の結果に関するもので、最終結果に対するものではない．デジタル計算機は、この「雑音レベル」の問題に関しては、アナログ計算機に対して非常に有利である．デジタル計算機における精度の問題は、アナログ機械とはまったく異なることから生じる．乗算を例にとると、二つの10桁の数の積は20桁になるが、ここにはまったく誤差はなく、精度は絶対的なものである．これはアナログ機械では決して望めないものであり、デジタル機械の大きな利点である．しかし実際にはデジタル計算機でも誤差が生じてしまうのである．というのは、たとえば二つの10桁の数の積は20桁であるが、計算機は10桁の数しか扱うように作られていないとすれば、20桁の数の下10桁は無視して、最初の10桁のみを扱わねばならない．このようにして生じる誤差は「丸め」(“round-off”) 誤差と呼ばれる．

アナログ計算機とデジタル計算機における誤差はまったく性質の異なるものであるが、両者を

*²³ Von Neumann, “The General and Logic Theory of Automata”, *Collected Works*, Vol. 5, p.289.

量的に比較すると後者の方がきわめて有利である。アナログ機械における「相対的雑音レベル」は $1:10^5$ より小さいことはなく、多くの場合において $1:10^2$ にもなる。一方デジタル機械では、10進法で10桁を扱う場合には「相対的雑音レベル」は $1:10^{10}$ となり、誤差のレベルは、何桁の数を設計上扱うかに依存している。以上の点から、以下の考察では計算機として議論の対象とされているのはデジタル計算機のみである。

第3章「計算機と生物体との比較」では、デジタル計算機と中枢神経系、とくに両者を構成する素子が比較され、ニューロンの持つアナログとデジタルの混合的特性に注目されている。ニューロンが伝達する信号は二値の性質を持つのでデジタル過程と見なすことができるが、一方で血液などの体液の化学的成分が生物体の振舞を決定していることも知られており、それはアナログ過程である。たとえば血圧を一定に保つ機構はデジタルとアナログの両方の機構からなっており、混合的な機構が重要な役割を果たしている場合も知られている。生物体ではアナログ的な要素が認められるが、考察に際しては、単純化のためにそれを捨象し、純粋なデジタル・オートマトンとみなすことにする。ニューロンに関しても、その混合的特性のために、その機能は非常に複雑であるが、考察に際してはニューロンは「全てが無かというデジタル器官であるという作業仮説」を用いることとされている。すなわちニューロンを電気的器官とみなし、その機能の化学的な側面は無視して、真空管と同様に「ブラックボックス」として扱っている。

このように捉えれば、ニューロンは「生物体の基本的スイッチ器官」であり、計算機の場合にその役割を果たすのは真空管や電気機械式リレーである。回路を構成する素子としてのスイッチという観点から、大規模計算機と生物体を比較すると、複雑さに関して前者が後者に及ばない理由が理解される。第一には構成素子の大きさである。真空管はニューロンの十億倍の大きさを持っており、生物体と同じ複雑さを有する回路を構成することは現在の技術ではまったく不十分である。また第二の理由として構成素子の信頼性が挙げられる。

これらの問題は素子に関する物理的なものであるが、人工オートマトンが自然のオートマトンに及ばない第三の理由は理論的な問題である。すなわち「オートマトンの論理的理論」が存在しないのである。これが第4章「オートマトンの将来の論理的理論」のテーマである。たしかに非常に精緻な形式論理学の体系は存在しており、数学にも適用されているが、しかしそれは実数や複素数といった連続的な概念にはほとんど関係せず、解析学には無縁なのである。解析学こそ現在最も技術的には成功した分野であるが、形式論理学はそのアプローチの性格によってこの分野から切り離されている。現在のオートマトン理論はデジタル的、二値的タイプのものであるので、形式論理学と同じ性格を持っている、すなわち数学的な視点からは解析学的というよりも組合せ論的なのである。

さらにオートマトン理論が抱えている問題には、形式論理学では存在しないものがある。論理学において重要なことは、ある結果が有限な数のステップによって達成可能か否かということであって、そのために必要なステップの数が問題となることはなかった。しかしオートマトン理論の場合には、どれだけのステップが必要かが重要な問題となるのである。その理由は二つであ

る。第一に、オートマトンは、前もって与えられた時間内に特定の結果に到達するように作られねばならない。第二に、各構成要素は、個々の演算においてわずかだが零ではない確率で誤動作をするかもしれないのである。個々の確率は小さくとも、演算数が多くなるに従って誤差が累積し全体の誤動作の確率は大きくなる。

以上の点を考慮すると、オートマトンの論理は、形式論理学の体系とは二つの点で異なることになる。

1. 「推論の連鎖」すなわち演算の連鎖、実際の長さが考慮されねばならない。
2. 論理演算(演繹法、連言、選言、否定など、すなわちオートマトンに関して一般的な用語では、様々な形態のゲート、一致、反一致、ブロッキングなどの動作)はすべて、低いが零ではない確率を持つ例外(誤動作)を許容する手続きでもって扱われねばならない^{*24}。

ノイマンによれば、このような要請を満たす理論は厳格で二値的な形式論理学からは遠ざかり、解析学的なものになるはずである。言い換えれば、さらに解析学的オートマトンそして情報の理論が必要なのである。

人間の中枢神経系のようなきわめて複雑なオートマトンを作る上で最も重大な制限要因は、誤りを処理する手続きがないことである。生物体では、構成要素の誤動作が生じると、それは検知されて、その影響が最も小さくなるようにシステムが調整され、全体の機能は保持されるようになっている。人工オートマトンの場合には、すべての誤動作をできるだけ早く検知し、問題となる構成要素をすぐに修理あるいは交換しなければならない。この過程は自動的に行なわれる場合もあるが、その診断の重要な部分は外部からの介入によって実行されねばならない。自然における誤動作を処理する際の基本原理とは、その結果をできるだけ重要でないものにすることであって、必要があるならば、余裕のあるときに訂正するのである。一方人工オートマトンの場合にはすぐに診断し調整しなければならず、そうしなければより状態が悪くなる可能性が高いだろう。またわれわれの誤りの診断法は、ただ一つしか異常のある要素はないことを前提としている。機械を繰り返し部分に分割していくことによって故障箇所を特定するのであるから、複数の異常な要素がある場合には、この方法は役に立たないのである。いかにして生物の持つ柔軟なシステムを人工物においても実現するのが、彼の課題だった。

第5章「デジタル化の諸原理」では、神経系において、連続的な量が二値のデジタル信号に変換される方法が論じられている。この符号化の機構として、展開法と計数法という二つの方法が述べられ、自然の組織は後者を採用しているとフォン・ノイマンは考えている。展開法とは数を桁によって表現し、各桁の数を送るものである。10進法で百万という数を伝えるには、各桁を表わす7個の数を表わすパルスを送ればよい。一方計数法は、その数分だけのパルスを送るものであり、たとえば百万という数を伝えるには百万個のパルスを送って数える。したがって展

^{*24} Ibid., p.304.

開法の方がはるかに計数法よりも効率がよいが、後者の方が安定しており、誤りにも強い性格を持っている。展開法においては一つのパルス伝達の誤りでも場合によっては大きな誤りを導くことがあるが、計数法ではそのようなことはない。この点からは自然では計数法が用いられていると考えるのが妥当だと判断しているが、もちろんそのような証拠が見いだされているわけではないことを注意している。

第6章「形式的ニューロン回路網」では「これまでに公理論的方法によって得られた最も重要な結果」*25であるマカロック・ピッツ・モデルが取り上げられる。このモデルについて概略を述べてから、二値しか持たないニューロンという仮定は「ニューロンの機能の過度の単純化」であることを認めている。フォン・ノイマンによれば、「彼らの試みの最も重要な成果とは、有限な数の言葉でもってまったく論理的に、厳密に、曖昧なく定義され得る、この意味での任意の機能が、そのような形式的ニューロン回路網によっても実現化可能である」*26ことを示したことにあたる。すなわち完璧に曖昧なく言葉にされうることは適切な有限のニューロン回路網によって実現されうるのであるが、しかし二つの問題が残されている。

1. ある種の行動様式が有限のニューロン回路網によって実行可能ならば、その回路網は実際のな規模内で実現可能か、問題となる組織の物理的な制限に適合するのか。
2. あらゆる存在する行動様式が本当に、完全に曖昧なく言葉にすることができるのか*27。

第一の問題は神経生理学の問題であるのに対して、第二の問題は論理学的含意を持つ問題である。すなわち二つの相似図形の認識といった活動は「完全に曖昧なく」言葉で記述することはできない。したがって高度に複雑なオートマトンとりわけ中枢神経系を理解するためには、本質的に論理的な新しい理論が必要である。

最終の第8章「複雑性の概念：自己複製」では、チューリングの理論を紹介し、彼の「万能オートマトン」理論に基づいて自己複製オートマトンの理論の構築することを試みている。複雑性に関して、自然と人工のオートマトンは異なる特性を持っている。人工オートマトンが作り出すものは自分より単純なものであるから、複雑性が減少するのに対し、自然界では、複雑性が減少しないものを複製し、長い進化の過程では複雑性は増大してきたことが知られている。この相違はどのようにして説明されるのだろうか。

そこでフォン・ノイマンは「チューリングの計算オートマトン」とくに「万能オートマトン」の理論を取り上げる。「万能オートマトン」とは、任意の他のオートマトンの機能を模倣するように組み立てられたオートマトンである。他のオートマトンが実行できる任意の演算を複製するためには、対象となるオートマトンの記述と、加えてその装置が考えられている演算のために

*25 Ibid., p.308.

*26 Ibid., p.309.

*27 Ibid., p.310.

必要となるはずの命令を与えることで十分である。チューリングのものは、その出力が0と1の記されたテープだから、それを拡張して出力がオートマトンであるようなオートマトンを作りたい。そこでフォン・ノイマンは、自己複製すなわち、一群の部品を一つの容器の中に入れたとき、最終的に最初のものと同じオートマトンであることがわかるもう一つの集まりを作り始めるようなオートマトンの青写真を提示している。それは、万能オートマトン、命令のコピーを作るオートマトン、それらを制御する機構から構成されたものであって、以下の5つのステップからなっている。

- (a) オートマトン A : 適切な関数によって他のオートマトンの記述が与えられたとき、それを構成する。その記述は命令 I によって示され、オートマトンはそれが挿入される場所を持つ。
- (b) オートマトン B : 与えられた命令 I のコピーを作る。
- (c) オートマトン A と B 、さらにそれらの制御機構 C を結合する。 C は A に、その命令 I によって記述されたオートマトンを構成させる。また C は B に命令 I をコピーさせ、そのコピーを構成されたオートマトンに挿入する。最後に C はこの構成物をシステム $A + B + C$ から分離し、独立した実体とする。
- (d) この集まり全体 $A + B + C$ を D によって示す。
- (e) 集まり $A + B + C$ は、機能するためには命令 I を与えられねばならない。オートマトン D を記述する命令 I_D は、 D 内の A に挿入されねばならない。こうして作られた集まりを E と呼ぶ。 E は自己複製的である^{*28}。

ここでの命令 I_D は遺伝子の機能を果たしており、複製機構 B は遺伝物質の複製機構に対応している。系 E とくに I_D の変更は突然変異に結び付くものと考えられる。他のものの構成は、 I_D を、オートマトン D に加えて F も記述する I_{D+F} によって置き換えるによって行なわれる。 D において、その中の A に I_{D+F} が挿入されたとき、 E_F と表わすと、それは E を複製するほかに F も構成する。ここで命令が生物の遺伝子に相当し、命令の変更が突然変異に結び付くと考えられる。

この論考における自己複製機構に関する記述は以上だけであって、抽象的な段階にとどまっていたが、「複雑なオートマトンの理論と構成」においてはさらに詳しく述べられていた。ここでは、その8種類の部品やそれらの接合方法などが論じられ、人間の組織を見本として物理的なイメージを伴う部品からの構成が論じられている。この論考の編集者であるパークスは、このモデルは「運動、接触、位置ぎめ、溶接、切断というような幾何学的運動学的な問題を取り扱い、力とエネルギーという本当に力学的、化学的な問題は無視した」ことから「運動学的モデル」(“kinematic model”)^{*29}と呼んでいる。

^{*28} Ibid., pp.316-317.

^{*29} *Theory of Self-Reproducing Automata*, p.82.[『自己増殖オートマトンの理論』, p.100.]

4. 5つのオートマトン・モデル

「オートマトンの一般的かつ論理的理論」と「複雑なオートマトンの理論と構成」では、中枢神経系を規範とし、生物と人工物における構成素子の相違を確認した上で、いかにして人工的に中枢神経系に近いオートマトンを可能にするような理論を構築することが問題とされ、その手がかりがマカロック・ピッツ・モデルとチューリングの理論に求められていた。生物のような自己複製能力は人工物においてもいかにして可能になるのか、またそれを可能にする理論の必要性が説かれていたが、その具体的な姿は展開されていなかった。その行き詰まりの原因は、彼が物理的に生物に似たシステムを構築しようと考えたことにありと云えよう。部品の接合方法といった、オートマトン理論とは本来関係のない問題が入ってきてしまい、そのために不必要にシステムが複雑になってしまったのである。フォン・ノイマンに自己複製オートマトン理論の探求の突破口を示唆したのは、ウラム (S. M. Ulam, 1909-1984) だった。ウラムは、彼がそれまで展開してきたモデルを放棄し、「細胞モデル」(あるいは「セル・モデル」“cellular model”) を研究することを勧めたのである^{*30}。このモデルは、それまでのモデルで問題とされていた物理的な素子の結合といった問題を回避し、純粋に数学的な検討だけで済むという利点を持っていた。この「細胞モデル」に基づくオートマトン理論は「オートマトンの理論：構成、複製、均質性」として未完のまま遺稿として残されたが、パークスによってまとめられ、『自己増殖オートマトンの理論』の第二部として出版されている。パークスは、自己増殖オートマトンとして、フォン・ノイマンは5つのモデルを考えたと言及し、以下のような名称を付けている。

- (1) 運動学的モデル (kinematic model)
- (2) 細胞モデル (cellular model)
- (3) 刺激 - 閾 - 疲労モデル (excitation-threshold-fatigue model)
- (4) 連続モデル (continuous model)
- (5) 確率的モデル (probabilistic model)^{*31}

第1のモデルは、「オートマトンの一般的かつ論理的理論」や「複雑なオートマトンの理論と構成」において自己複製や自己増殖の問題を扱った際に言及されたものであって、機械的な装置としてのオートマトンの可能性を論じたものだった。第2のモデルはそれら論考には現われないが、先ほど触れたようにウラムの示唆を受けて考察が始められたもので、この論考の考察の中心となっている。第3以降のモデルは第2のモデルを基礎として考えられているが、実際にフォン・ノイマンが具体的な理論を生前に展開できたのは第2のモデルだけだった。

^{*30} Ibid, p.94.[同上, pp.112-3.]

^{*31} Ibid., p.93.[同上, p.112.] フォン・ノイマン自身も書簡において「細胞モデル」を「結晶モデル」と呼んでおり、「刺激 - 閾 - 疲労の機構をもったモデル」、「連続モデル」という表現も用いている (Ibid., p.95.[同上, p.114.]).

「細胞モデル」^{*32}は、同一の有限オートマトンが一つずつ入った細胞が無限に広い平面上に並んでおり、各オートマトンは 29 の状態を持っている。各細胞は周囲の細胞と直接に 1 単位時間あるいはそれ以上の遅れを持って通信し合いながら自己増殖を行なっていく。第 3 の「刺激 - 閾 - 疲労モデル」では、「細胞モデル」における 29 状態を持つオートマトンを、閾と疲労の機構を持つ神経細胞的な素子から構成するというものである。これによって、より現実の神経回路網に近いようなモデルを作ることを目指した。第 4 の「連続モデル」は、それまでのモデルが数理論理学に基づく離散的なモデルだったのに対して、信号は連続的に変化するものとして、それを連立偏微分方程式によって扱うというものである。実際の細胞における興奮、閾、疲労といった現象はアナログ的なものであるから、その記述は偏微分方程式によって扱われるのが適切なのである。自己増殖モデルは離散的なものから連続的なものへと進むが、それは人工的な機構から自然の機構へ進もうという彼の研究目標を表わしていた。ちょうどこれは、彼が衝撃波といった自然現象を扱う際の過程とは反対であった。衝撃波のような非線形な現象では、衝撃波の進行する液体を小空間に分割し、その中では一様とするモデルを作る。それによって連続な系の微分方程式は離散系の非線形な差分方程式に置き換えられ、その数値解をデジタル計算機によって求めることによって、もとの微分方程式の近似解を得るのである。

こうしてフォン・ノイマンは「細胞モデル」においても、より自然のものに近いような自己増殖オートマトンを求めていったが、同じ意図は第 5 の「確率的モデル」にも見いだされる。このモデルについては、この論考の中ではほとんど述べられていないが、それは、セル・オートマトンにおける状態の遷移が、そのセルおよび周囲のセルの以前の状態によって一意的に決定されるのではなく、確率的に決定されるというものである。これによってオートマトンにおいて、突然変異や進化という現象をモデル化しようとしたのだった。より簡単なものからより複雑なものに、効率の悪いものから効率のよいものにどのようにして進むのかをこのモデルの中で探求しようというのである。

5. オートマトン・モデル 人工から自然へ

フォン・ノイマンのオートマトン理論の出発点は、マカロック - ピッツ・モデルが示唆している、人間の中枢神経系とコンピュータの情報処理が基本的には同一のものであるという視点だった。ニューロンの生理的な性質を捨象し、入出力の機能のみが規定された「ブラックボックス」として扱う「公理的アプローチ」は、全体の機能がどのようにして個々の素子から実現されるのかという問題に焦点を定めることにおいて実り多きものだった。

電子計算機において、非常の多くの素子から構成される回路を考えると、その複雑さに関して模範となり得たのは神経回路だけだった。しかし形式論理学によって扱うことができたのは、理想化された神経細胞すなわち「形式的ニューロン」であって、現実のものではなかった。実際のニューロンはアナログ的な側面も持ち、また神経系は複雑性に関しては、人工のものとは比較

^{*32} 我々には「セル・オートマトン」という言葉の方が馴染みがある。

にならないレベルにあった。さらに信頼度の低い素子から成るシステムが高い信頼性を持っているという特徴を備えていた。

人工オートマトンのハードなシステムを、いかにして自然オートマトンのソフトなシステムへと発展させるかがフォン・ノイマンの課題だった。そこで彼が考えていたことは、デジタルからアナログへ、組合せ論から解析学へと進むことである。現在のオートマトン理論が抱える問題の大多数は数理論理学に起因するものであるから、それを離れて非線形偏微分方程式を用いて解析学に近づくことが求められた。このとき彼の頭の中には、通信理論において情報量の概念が熱力学のエントロピー概念によって解析的に扱われていたことがあった。

「オートマトンの一般的かつ論理的理論」では、このオートマトン理論の解析学化について次のように述べられていた。

形式論理学を研究している人は誰もが、それは数学のなかでも技術的に最も厄介な分野であることを認めるだろう。この理由は、それが、厳格ですべてが無かという概念を扱い、実数や複素数といった連続的な概念すなわち数学的解析とほとんど接触がないからである。だが解析学は数学のなかで技術的に最も成功し、最も精緻な分野である。したがって形式論理学は、その手法の性質によって、数学のなかで最も洗練された部分から切り離されており、組合せ論という数学という領域のなかでも最も困難な部分に押し込まれているのである^{*33}。

これらすべてのことは、先に指摘された結論、すなわちオートマトンと情報に関する、詳細で、高度に数学的で、いっそうとりわけ解析的な理論が必要であるという結論を改めて強調している^{*34}。

フォン・ノイマンは、オートマトンのためには組合せ論よりも解析学に近い理論が必要と感じていたが、そのような理論はまた確率論的および統計学的性格も持つものだろう。それは、信号処理上の誤りを自動的に検査し修正するシステムを作るためにも不可欠だった。彼のオートマトン理論はデジタル理論を出発点としながら、高度に複雑なシステムとして神経系を模範とすることによって、アナログ的、解析学的な側面を探索していくものだった。彼は自然と人工の世界を統合するようなオートマトン理論を求めたが、それは実現することはなかった。しかし彼にとって、すべての研究は自然を出発点としており、彼の作るモデルはすべて自然に対するモデルだったのである。自然と人工のオートマトンの統合というフォン・ノイマンの試みは、彼の死によって中断され、それを受け継ぐ者も現われなかった。彼の試みはきわめて特異なものに思われるが、信頼性の低い素子からなるが、全体としては信頼性の高いシステムをどのようにして構築するかという試みは現代のわれわれにとっても示唆に富むものである。

^{*33} “The General and Logic Theory of Automata”, p.303.

^{*34} Ibid., p.304.