

# 人工生命からシステム生物学へ

大塚淳<sup>1</sup>

## 1. はじめに

人工生命を主題とする世界最初の国際会議がニューメキシコ州ロスアラモス国立研究所で開かれたのが、1987年。それから約20年がたった現在、コンピュータシミュレーションを用いた生物学的研究は、すでに一定の市民権を得ているとっていいだろう。これには、ロバート・ラングトン、ジョン・ホートン・コンウェイ、スチュアート・カウフマンなどの人工生命研究の「スタープレイヤー」たちの活躍がある。こうしたこともあり、人工生命、あるいは生命のシミュレーションというと、スクリーンの中で点滅を繰り返す「セル」たちや、色とりどりのアトラクター、「創発」や「カオスの縁」などといったキャッチーなキーワードが思い起こされるのが一般的なのではないだろうか。

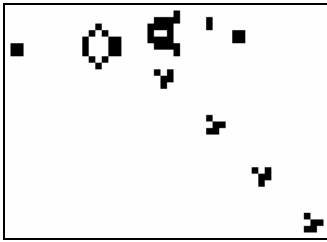
しかしこのような研究の一方で、今世紀に入ってから、システム生物学という新たな研究分野が開拓されてきた。この潮流に属する研究は、生物学的研究にコンピュータシミュレーションを用いるという共通点は有するものの、その歴史的経緯や方法論、さらには研究の目的などにおいて、従来の（あるいは一般的に理解される）人工生命研究とは毛色が異なっている。つまり人工生命とシステム生物学では、コンピュータシミュレーションに託された認識論的な意味づけが大きく異なるのである。しかしそれはどのような点においてであろうか。そしてその違いは両者の研究目的に対しどのような意味をもつのだろうか。本稿ではこの両者の比較を通して、生物学研究においてシミュレーションが担う役割を考察してみたい。

## 2. 人工生命

### ■ 人工生命とは

人工生命についてはすでに一般的な紹介文献が多数あるため(Emmeche, 1994; Keller, 2002; Levy, 1992)、詳しくはそれらを参照していただくことにして、ここでは以下での考察に必要なだけの概略を示すだけにとどめよう。

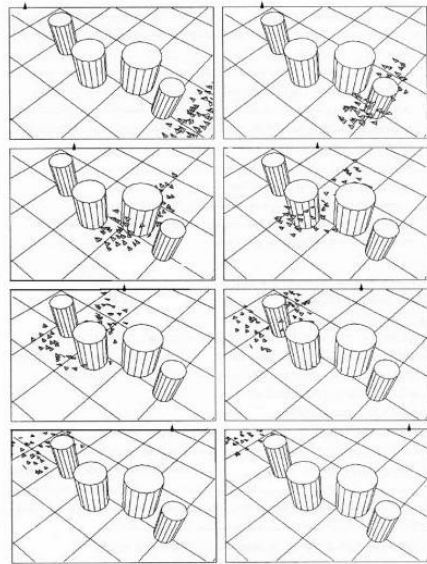
コンピュータを用いた人工生命研究の発端となったのは、フォン・ノイマンのセル・オートマトンである。良く知られているように、彼の動機は自己複製する機械を作り出すことであり、セル・オートマトンはそれを実現する形式的な媒体であった。有名なコンウェイ



(図1) ライフ・ゲーム。グライダーを打ち出すグライダーガン

イの「ライフ・ゲーム」は、このフォン・ノイマンのセル・オートマトンを単純化したものである(図 1)。ライフ・ゲームにおいては、各セルの状態は2つに限定され、それは隣接する8つのセルの状態に応じて変化する関数になっている。このような単純な規則によって支配されるセルを二次元空間に碁盤の目状に敷き詰め、適当な初期状態からスタートさせると、セルたちは様々な遷移パターンを示していくのである。他にも良く知られている

人工生命の一例として、レイノルズの「ボイド (boid)」がある(図 2)。各ボイドは、次のような三つの簡単な規則に従って仮想空間を移動する。(1)周囲のほかの対象(ボイドや障害物)から、最小限の距離を取る；(2)近隣のボイドと速度をそろえて進む；(3)近隣のボイド個体群の中心に向かって進む。各ボイドの挙動を支配するのはこうした一見単純なルールにすぎなくても、これらボイドからなる集団は、鳥などの群れが示す均整かつ同期の取れた挙動を非常にうまく再現するのである。また比較的近年の研究としては、後に取り上げる、遺伝子の相互作用をブール代数のネットワークで表現したカウフマンのNKモデルが挙げられるだろう。



(図2) レイノルズのボイド。群れは円筒形の障害物を避けながら進む。ボイド自身によるアプレットが <http://www.red3d.com/cwr/boids/> にある。

## ■ 人工生命の目的と方法論

人工生命研究は、このように多岐な形態をとっているが、底に流れるテーマは共通している。共通項は、(1)単純な部分と規則の組み合わせが、全体における複雑なパターンを作り出すという発見と、(2)そうして作り出されるパターンが、生命の重要な特徴を捉えているという主張である。前者は一般に、創発(emergence)と呼ばれる。後者の主張は、人工生命研究者が掲げる、全体論的(holistic)なアプローチを要約している。つまりそれは、部分的な振舞いよりも、全体的なパターンが説明において果たす役割を重視し、そうした全体の性質によってこそ生命の理解が得られると主張する点において、全体論的だと言えるのである。

では、このようなアプローチから得られる知見とは何だろうか。単純な規則から複雑なパターンが生じるという発見は、確かに驚きに値する。さらにそれが現実の生物のふるまいと似ているとしたら、驚きはなおさらだろう。しかしそこから、生命に関するどのようなことが「分かった」ことになるのだろうか。人工生命研究者が主張するその認識論的意義は、**生命に普遍的な法則**の発見である。例えばラングトン<sup>1</sup>は以下のように述べる：

理論生物学は長い間、単一の事例からは一般的な原理を導くことができないという根本的な障害に直面してきた。他の事例が見つからない限り、生命の本質的性質——原理的にあらゆる生命システムに共有されていなければならない性質——を、単に局所的な歴史的偶然の組み合わせと遺伝的相続の共通性<sup>2</sup>だけからたまたま地球上生命において普遍的となっているような、生命にとって偶有的かもしれない性質から区別することは極めて困難である。この先の探求において、異なった物理化学に基づいた生命体が我々の前に現れてくるようなことはとてもありそうにない。したがって我々の唯一の選択肢は新たな生命形式、すなわち人工生命の合成を試みることである。(Langton, 1996, p. 39, 強調はラングトン)

地球上の生命は多様であり、その多様性は偶然的要因（例えば突然変異や環境変動）によって支配されている。グールド<sup>3</sup>が言うように、「生命のテープをリプレイ」<sup>4</sup>するとしたら、その結果は現在の我々が見知っている生物界とは大きく異なったものになるだろう(Gould, 1990)。しかし何回そのような「リプレイ」を繰り返したとしても、その全ての登場人物（あるいは登場生命）が必ず備えているような普遍的な性質、すなわち生命の本質はないのだろうか。もしそのような性質を見つけられれば、我々は、一見多様で無規則的な生物界において、実はすべての生物が従っているような普遍的法則を手にするようになるだろう。

カウフマン<sup>5</sup>のシミュレーションは、こうした生命の普遍的法則を探ろうとするものである(Kauffman, 1992, chap. 12)。彼は自らのNKモデル<sup>6</sup>を、ゲノム内の各遺伝子が調節タンパクを通して互いの発現に影響を及ぼしあうネットワークのモデルとして解釈し、その上でこのモデルを構成するノードの数と、モデルが安定した挙動を示すときのパターンの周期長とを比較した。その上で、各々ノード数は一定だがノードの間の配線をランダムに組み替えた多数のNKモデルをコンピュータ上で走らせその周期を測ると、ノード間配線に関わらず、ノード数と周期長には一定の関係があることが発見された。彼のモデルの解釈に従えば、これはゲノム内の遺伝子の数と細胞周期の間に一定の関係があることを示唆する。そしてカウフマンは、遺伝子数と細胞周期のこの関係は、現実の生物においても見出されると主張する。つまり、各生物種においてゲノムネットワークの具体的な配線は

様々に異なっていたとしても、そのような差異に関わらない普遍的な関係性が、あらゆる生物の遺伝子発現において成立しているというのである。

このように見ていくと、人工生命研究の目標（普遍法則の探求）と、本節冒頭で掲げたその研究アプローチ（全体論）は、実は密接に関係したものだということが理解される。すべての生命——人工生命研究者がこういうとき、そこには地球上に存在する／した生物だけでなく、存在しえたらろう可能的な生命も含まれる——に共通する性質を特定するためには、個々の具体的な生物が持つ詳細な情報はできる限り捨象されなければならない。そうした詳細は、極めて少数の本質的な要素を覆い隠すノイズでしかない。上述のカウフマンが主張する法則は、個別的な遺伝子「配線」についての情報を捨象し、遺伝子数と細胞周期というたった2つのパラメータに注目したときにはじめて浮かび上がってくるものである。このように人工生命研究にとって、データ量のコンパクトさは、制約ではなくむしろ要請なのである。一方、次の節で見るように、システム生物学における方法論は、人工生命研究のこうした全体論的なアプローチとは、好対照をなしているのである。

### 3. システム生物学

#### ■ システム生物学とは

システム生物学とは、その名の通り、生物を一つのシステムとして理解していこうという学問領域である。生物を構成する要素の機能単体ではなく、それらの組み合わせに着目するというアイデア自体は、新しいものではない<sup>2</sup>。しかしその研究が具体的なデータに基づく形で現実味を帯び始めたのは、この十年間においてである(Kitano, 2001)。これには第一に、ゲノミクスやプロテオミクスなどの興隆に見られるように、生命機構に関する分子レベルでの莫大なデータが、近年急速に蓄積されてきたということがある。そしてこの背景には、DNAマイクロアレイなどといった、高速かつ大規模な測定を可能にする技術の進展がある。そしてもちろん、大量のデータを管理・処理するためのコンピュータパワーの向上も、こうした研究の必要条件としてあっただろう。

このような流れを受けてスタートしたシステム生物学の目標の一つに、様々な生命現象、たとえば細胞内の代謝活動や免疫反応、多細胞生物の胚発生などを、コンピュータ上(*in silico*)に分子レベルでシミュレートするというものがある。日本における研究としては、ERATO-SORST 北野共生システムプロジェクト (<http://www.symbio.jst.go.jp/symbio2/index.html>)、慶応大学の富田勝を中心にした E-CELL Project(<http://www.e-cell.org/ecell/>)、また国外では Alliance for Cellular Signaling (<http://www.afcs.org/aboutus/index.html>)、Virtual Cell Project(<http://www.ncam.uchc.edu/>)、Silicon Cell Project(<http://www.bio.vu.nl/hwconf/Silicon/>)

などの例がある。

これらのプロジェクトから具体的な研究例を少し見てみよう。北野らは、ショウジョウバエの初期胚に見られる分節化現象をシミュレートしている(Hamahashi & H. Kitano, 1998; 北野, 2001b, 2001a)。ショウジョウバエ胚発生の初期に現れる体節は、複数の遺伝子産物(タンパク質)が胚の前後軸に沿って形成する濃度勾配のパターンによって決定される。そしてこのパターンは、母性効果遺伝子(*bicoid*)が引き金となって起こる、一連の制御と遺伝子発現パターンによってもたらされる。北野と濱橋の研究は、遺伝子の転写速度、タンパク質の拡散係数や分解速度などを定めた上で、この発現パターンをシミュレートし、実際のショウジョウバエで観察される遺伝子産物の濃度パターンが得られるかどうかを探るものである。

一方、富田らは、彼らの開発した E-CELL システムを用いて、ヒトの赤血球細胞をモデル化している(富田, 2001)。このモデルはおよそ 80 種類の異なる要素(代謝中間体および酵素)を有しており、解糖系・ペントースリン酸経路・核酸合成経路という 3 つの代謝経路の働きをシミュレートする。彼らのシステムでは、こうした酵素のうち特定のものを欠損させたとき、それが他の要素の状態にどのような影響を与えるのかを、確認することができるようになっている。

## ■ AL との違い

このような概観からだけでも、システム生物学と前節で扱った人工生命研究との間の差異が見て取れるだろう。第一の違いは、人工生命が要素の単純性あるいは等質性を仮定していたのに対し、システム生物学においては、システムを構成する諸要素にそれぞれに異なった役割が与えられるということである(Cf. Kitano, 2002)。

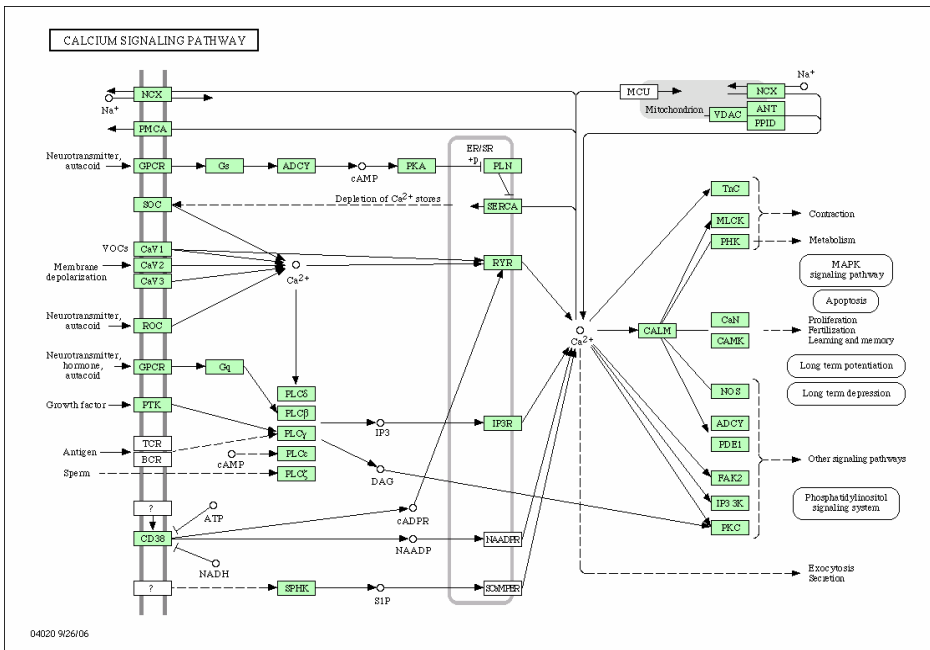
また後者の研究では、シミュレート対象である系の細部に関する詳細な情報が本質的な重要性を帯びてくる。濱橋らのシミュレーションでは、遺伝子の初期分布濃度やタンパク質の拡散係数の取り方によっては、まったく非現実的な結果が得られてしまう。このように、一般にシステム生物学では、人工生命研究は言わずもがな、一般的な生物学的研究と比べても、より広範かつ精確な定量的データが必要とされるのである(Kitano, 2001, p. 5)。

システム生物学と人工生命では、その目標も大きく違う。ここで目指されているのは、生命一般に妥当するような普遍的法則ではない。むしろシステム生物学におけるシミュレーションの目標は、詳細なデータに基づき、ハエの発生や特定の細胞の挙動などといった個別的现象をできるだけ忠実に再現することである。つまりここでは、普遍性ではなく、個別・特殊性こそが優先されているのである。

■ システム生物学におけるシミュレーションの意義

上述のように、システム生物学が目指すのは、普遍法則の発見ではない。むしろその目的は、個別的な生命現象のもとにある**因果関係を明らかにすること**である。つまりシステム生物学におけるシミュレーションの意義は、因果モデルとしてのその役割にある。

現在一般的な（かつシミュレーションとは異なる）因果モデルとして、パスウェイ解析があげられるだろう。一般にパスウェイ解析では、システムの構成要素の機能的特徴とその機能連関が、それぞれノードと矢印によって表される。このような表現は分子生物学でも一般的であり、例えば京都大学で進められている生命システム情報統合データベース KEGG(<http://www.genome.jp/kegg/>)では、代謝や遺伝子発現など、様々な細胞機能に関する反応経路モデルが、こうしたダイアグラムの形で網羅的に収められている（図3）。



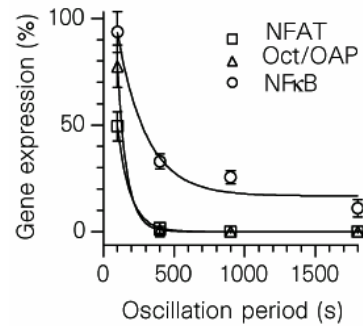
(図3) KEGG PATHWAY より。図はヒトのカルシウムシグナル経路。四角はタンパク質、丸は代謝中間体、矢印は各要素間の因果関係を表している。

ではこのように人口に膾炙した因果モデルがあるのにも関わらず、なぜ新たな因果モデルとしてシミュレーションを持ちだす必要があるのだろうか。その答えは、対象の通時的変化にある。パスウェイ解析においては基本的に、各要素と関係は固定されている。つまり時間の進展にともなって、それらの機能や結びつきが変化していくことは想定されていない。しかしながら生命活動においては、構成要素の役割やその間の関係の通時的な変化はきわめて一般的な現象である。

例えば、カルシウム( $\text{Ca}^{2+}$ )は細胞内での様々な応答を引き起こす重要なシグナル分子であるが、これによって引き起こされる反応のうちには、その濃度変化の周期に依存するものがある。通常、細胞内の $\text{Ca}^{2+}$ 濃度は低く抑えられているが、一定の刺激を受けると、その濃度は急激な上昇と下降のスパイクを示す。そしてこのスパイクの振動数に応じて、細胞は異なった仕方で反応する。例えば低振動数では遺伝子Aが、高振動数では遺伝子A

およびBが発現されるというように、 $\text{Ca}^{2+}$ シグナルの帰結はその濃度周期によって決定されるのである(図4)。したがって、 $\text{Ca}^{2+}$ シグナルの役割を知るためには、ただそれがどのような経路に組み込まれているかだけでなく、どのようなときにどの経路が活性化されるのか、ということも知らなければならない。そしてその条件となるのは振動数という時間的要素である。つまり、 $\text{Ca}^{2+}$ シグナルを組み込む経路の挙動を完全に理解するためには、時間的契機(これには単位時間ごとの $\text{Ca}^{2+}$ の濃度変化などといったパラメータが含まれるだろう)を持つモデルが必要となってくるのである。

シミュレーションの利点は、このような時間的変化を扱うのに適しているということである。例えばMichelson & Schulman (1994)は、 $\text{Ca}^{2+}$ シグナルを媒介するCaM-Kinase ( $\text{Ca}^{2+}$ /カルモジュリン依存タンパクキナーゼ)が、 $\text{Ca}^{2+}$ 濃度のスパイクに応じてどのような時間的发展を示すかをシミュレートすることで、前述の振動数依存的な $\text{Ca}^{2+}$ シグナルの役割をモデル化している<sup>3</sup>。彼らのモデルは単にCaM-Kinaseの状態がどのように変化していくかを示しただけであったが、これをさらにCaM-Kinaseの持つ特異的な転写因子発現機能と結びつけたモデルを作ることも可能であろう。そしてそのようなモデルは、 $\text{Ca}^{2+}$ が細胞内において単にどのような役割を担っているのかだけでなく、いつ・どのような状況でそれら



(図4) T細胞において、 $\text{Ca}^{2+}$ シグナルはNF- $\kappa$ B, NF-AT, Oct/OAPの三つの転写因子をコントロールしている。低振動数ではNF- $\kappa$ Bのみが活性化されるのに対し、高振動数ではそれに加えNF-AT, Oct/OAPも活性化される(Dolmetsch, Xu, & Lewis (1998)より)。

の役割を果たすのかについても教えてくれるであろう。

このように、シミュレーションの利点はモデルに時間軸を組み込むことで、因果モデル自体の通時的変化を表すことができるということにある。さらにこのことにより、ある一つの経路の働きが、他の経路の働きとどのように同期し、前者における変化が後者にどのような影響をもたらすのかを示すことができる。細胞内では、一つの要素が、複数の経路において様々な役割を果たしていることが一般的である。こうした多機能的な要素を通じて、複数の経路が複雑に絡み合い、相互に影響を与え合っている。このような状況のモデル化を考えると、パスウェイ解析のような因果ダイアグラムでは、ある要素が複数の経路に組み込まれていることは（各経路のダイアグラムにおいて同じ要素を繰り返し登場させることで）理解されても、それによって各経路がどのように連絡しあうのかを表すことはできない。一方シミュレーションモデルは、時間という共通通貨を与えることで、これらの異なった経路の相互連関をモデル化することができるのである。

また、こうしたシミュレーションの持つ認識論的意義は、これらを用いた研究が詳細かつ精密な定量的情報を要求するという点とも符合している。というのもこうしたモデルにおいては、一つの経路に組み込まれたパラメータにおける若干の誤差が、他の経路に対し甚大な影響を及ぼすこともありうるからである。つまり、たった一部分における小さな誤りが、シミュレーションのモデルとしての価値を台無しにしてしまう。そしてこれを避けるためには、各パラメータを精密に設定するしかない。こうして、システム生物学におけるシミュレーションと人工生命研究のそれとの間の方法論的な違いは、両者の目的の違いと連動しているのである。

#### 4. まとめ

人工生命研究の父、フォン・ノイマンは、セル・オートマトンを提案するにあたって、こうした形式化が、問題の半分、しかもより重要かもしれない半分を投げ出すことに等しいのではないかと自問している(Von Neumann, 1966, p. 77)。捨てられた半分とは、彼が生命活動の本質と見なすものが、いかにして現実の物質的基盤から生じてくるのかという問題である。その後の人工生命研究は、フォン・ノイマンのアプローチに従いながら、ひたすら生命の形式的な本質と法則を追い求めてきた。今世紀の変わり目に産声をあげたシステム生物学は、実際の生命の分子的基盤と、それらの分子が織り成す因果関係のネットワークをひとつひとつ特定していくことで、半世紀前に捨てられ等閑視されてきた問題のもう「半分」を、地道に拾い集めていこうとする試みだと言えるのではないだろうか。



## 註

<sup>1</sup> junotk@gmail.com

<sup>2</sup> 例えば北野 (2001a)は、先行研究として、ベルタランフィーやウィーナーなどの研究を挙げている。

<sup>3</sup> このシミュレーション結果については、De Koninck & Schulman (1998)がin vitroでの「追試」を行っている。

## 文献

- De Koninck, P. and H. Schulman. (1998). "Sensitivity of CaM kinase II to the frequency of Ca<sup>2+</sup> oscillations", *Science* 279:227-230.
- Dolmetsch, Ricardo E., Keli Xu, and Richard S. Lewis. (1998). "Calcium oscillations increase the efficiency and specificity of gene expression", *Nature* 392(6679):933-936.
- Emmeche, C. (1994). *The Garden in the Machine: The Emerging Science of Artificial Life*. Princeton University Press. (1998, 佐倉統・神成淳司・高木和夫・山本貴光訳, 『マシンの園—人工生命叙説—』, 産業図書.)
- Gould, S. J. (1990). *Wonderful Life: The Burgess Shale and the Nature of History*. WW Norton & Company. (2000, 渡辺政隆訳, 『ワンダフル・ライフ—バージェス頁岩と生物進化の物語』, 早川書房.)
- Hamahashi, S. and H. Kitano. (1998). "Simulation of Drosophila embryogenesis", *Proceedings of the sixth international conference on Artificial life table of contents*:151-160.
- Kauffman, S. A. (1992). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.
- Keller, E. F. (2002). *Making Sense of Life*. Harvard University Press.
- Kitano, H. (2001). *Foundations of systems biology*. MIT Press Cambridge.
- 北野 宏明. (2001a). 『システムバイオロジー—生命をシステムとして理解する』, 秀潤社.
- (2001b). 『システムバイオロジーの展開—生物学の新しいアプローチ』, シュプリンガー・フェアラーク東京.
- Kitano, H. (2002). "Computational systems biology", *Nature* 420(6912):206-210.
- Langton, C. G. (1996). "Artificial Life." Pp. 39-94 in *The Philosophy of Artificial Life*. Oxford University Press.
- Levy, S. (1992). *Artificial life: the quest for a new creation*. Random House Inc. (1996, 服部桂訳, 『人工生命—デジタル生物の創造者たち』, 朝日新聞社.)
- Michelson, S. and H. Schulman. (1994). "CaM Kinase: A Model for its Activation AND Dynamics", *Journal of theoretical biology* 171(3):281-290.
- 富田 勝. (2001). 「E-CELL: ゲノム情報に基づく細胞の再構築」, 『生命情報科学の拡がり』, 102-108, サイエンス社.
- Von Neumann, J. (1966). *Theory of Self-Reproducing Automata*, edited by A. W. Burks. University of Illinois Press.

[哲学博士課程・日本学術振興会特別研究員]