

ブーリアンネットワークの固定点数の期待値の定理

森史¹, 望月敦史^{1,2}

¹ 理化学研究所 理論生物学研究室

² 理化学研究所 iTHEMS

概要

ON と OFF の 2 状態を取り得る要素が複雑に相互作用している系、例えば、遺伝子制御、神経回路、意見形成過程などのネットワーク系は、ブーリアンネットワーク (BN) で近似的に記述される。BN が広い状態空間の中でどのようなダイナミクスを示すのか、また、そのダイナミクスの特徴は、ネットワークトポロジーに対してどのような依存性をもつのか、という問題に対し、数学的な答えを用意することは、それらの系を理解する上で重要である。本稿では、任意のネットワークトポロジーを持つ BN において、各ノードのブール関数がある確率分布に従ってランダムに与えたときの、固定点数の数の期待値の定理について、解説する。

Theorem of the expected number of fixed points in Boolean networks

Fumito Mori¹ and Atsushi Mochizuki²

¹ Theoretical biology laboratory RIKEN

^{1,2} iTHEMS RIKEN

Abstract

Complex network systems consisting of elements that take either ON- or OFF-state, such as networks of genetic regulations, neurons, and opinion formations, can be approximately described by Boolean network (BN) models. In order to understand these complex network systems, we should clarify how the BN systems behave in a large state space and how their dynamics depend on the network topology. This brief note explains a theorem of the expected number of fixed points in a Boolean network with any topology and Boolean functions drawn from probability distributions.

1 はじめに

今夜の飲み会に参加するかどうかを、自分と気の合う A さんと、全く反りの合わない B さんの行動 (x_A, x_B) を参照した上で決めるとする。A さんが参加 $(x_A = 1)$ で、かつ、B さんが不参加 $(x_B = 0)$ のときのみ、自分は参加しようという意思決定の手続きは、ブール関数を用いて $f(x_A, x_B) = x_A \cdot \bar{x}_B$ と表現できる。ここで、 \bar{x} は、 x の否定を表す。A さん、B さんを含め、他の人もまた、それぞれのブー

ル関数を持っている状況で、誰一人不満を持たないよう参加者と不参加者を組分けする問題は、ブーリアンネットワーク (BN) の固定点を探す問題に他ならない (図 1)。

遺伝子制御ネットワークやニューラルネットワークでは、要素 (遺伝子、ニューロン) が、周りのオン-オフの状態を参照した上で、次に取る状態を決めていると捉えることができるので、BN によるダイナミクスの記述は妥当なモデル化と言える。BN は、

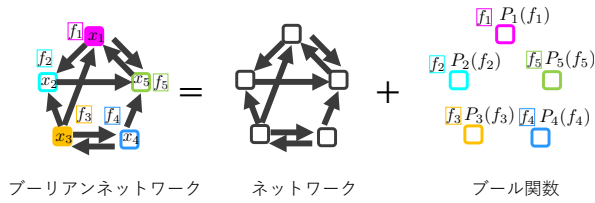


図 1: ブーリアンネットワーク (BN) の各ノードは、それぞれにブール関数を持つ。本研究では、ネットワークポロジを任意に固定し、ブール関数 f_i を確率分布 $P_i(f_i)$ に従ってランダムに割り振るとしたときの、固定点の数の期待値を考える。

要素の個性 (遺伝子の種類など) をブール関数の違いとして反映させることができ、各要素が 2 状態しか取らないにも関わらず、全体では多様なダイナミクスを示す。

BN の最も基本的な量の 1 つは、固定点の数である (図 2)。先に述べたように、BN を意思決定過程のモデルとして考えた場合、固定点はすべての人が意見を変えない状態であり、この意味で、すべての人が納得する解である。飲み会に限らず、政党の立ち上げや協定の締結へ至るプロセスを思い浮かべたとき、そもそもそのような解が存在しないのであれば、それを探す努力は無駄である。また、BN を細胞分化過程で働く遺伝子制御ネットワークとして捉えた場合、固定点は細胞の種類を決定する遺伝子発現パターンに対応すると考えられ¹、系は少なくとも細胞種類の数くらいには、固定点をもつ必要がある。このように、固定点解が存在するの否か、さらには、存在するのならいくつ存在するののかという問題は、様々な局面で重要である。

望月は、ランダムに構築した BN の平均固定点数が、系のサイズに依らずわずかに 1 であることを示し、このことから実際の遺伝子制御系はネットワークや制御関数に特別な構造を持っているべきであると指摘した [1]。また、スモールワールドやスケールフリーなど、特徴的なネットワーク構造をもつ系が自然界に多数存在することが知られて久しいが、ブール関数の選び方に全くバイアスをかけない BN では、平均固定点数はネットワークポロジに依存しないことも知られている [2]。一方で、いくつかの種類のバイアスは、摂動に対する応答などのダイナミクスの特徴を大きく変え、カオス-秩序転移を引き起こす [3]。そのようなバイアスは、固定点の数も

¹ 固定点に限定せずに、振動解まで含めたアトラクターとみなすべきという考え方もある。

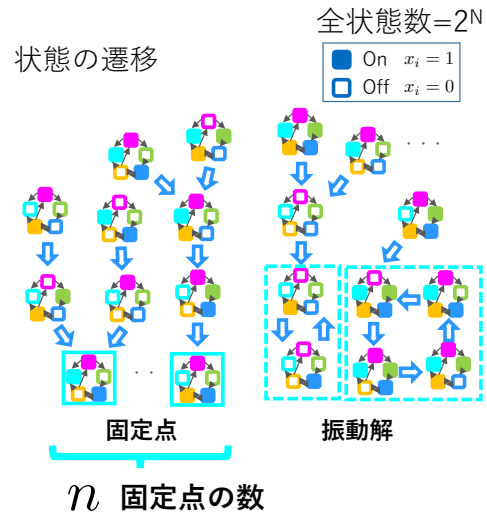


図 2: BN の状態遷移図。要素数 $N = 5$ として、模式的に描いた。すべての要素の状態を一斉更新するという取り決めの下では、図のような決定論的な状態遷移図が得られ、初期状態は固定点アトラクターか振動アトラクターのどちらかに行き着く。状態更新の順序についての取り決めを変えると、状態遷移図は一般に変化するが、固定点は変わらない。

変えるだろうか。また、何らかのバイアスを導入した上で、BN のネットワークポロジを変えると、固定点数は変化するのだろうか。文献 [4] では、その問いに答える定理を導出した。

2 ブーリアンネットワーク

2.1 モデル

BN において、ノード i の状態 x_i の状態更新は、

$$x_i(t+1) = f_i[\{x_j(t); j \in J_i\}], \quad (1)$$

で表される。 J_i は i への入力ノードの集合、 f_i は $\{x_j(t); j \in J_i\}$ を入力変数とする i のブール関数である。状態変数 x_i は、 $x_i = 0$ か $x_i = 1$ を取る。BN の時間発展を考える上では、状態更新をすべての i に対して一斉に行うか、あるいは、別々に行うかという取り決めが必要である。しかし、固定点はそれ以上は変化しない状態として定義されるので、その取り決めには依存しない (図 2)。

BN の固定点数に対するネットワークポロジの影響を調べるために、ネットワークポロジを固定した上で、各ノードのブール関数をそれぞれの確率分布 $P_i(f_i)$ に従ってランダムに割り振ることを考えよう (図 1、3)。図 3 は確率分布の例で、その横

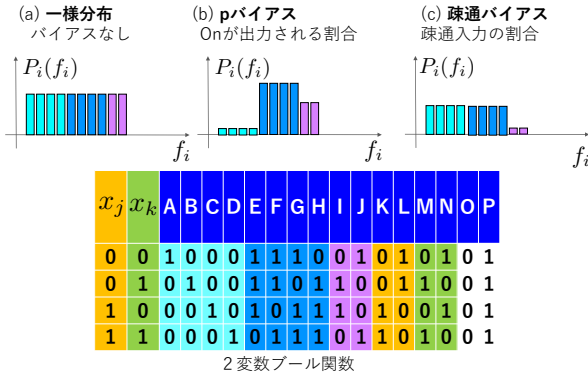


図 3: ブール関数の確率分布 P_i の例。グラフの横軸は、ブール関数 f_i に対応する。例えば、2変数関数の場合は、表のように 16 通りの関数が考えられる。図中では、すべての入力変数に依存する関数だけを採用するという流儀に従い、6 つの関数 (K-P) は採用していない (採用しても構わない)。(a) 一様分布。(b) p バイアス: ON を出力する確率 p を導入して、ランダムにブール関数を構築する。例えば、関数 A は、4 つの入力状態に対し、ON を 1 回、OFF を 3 回出力するので、出現確率は $P_i(A) \propto p^1(1-p)^3$ となる。等号でないのは、関数 K-P を採用しない場合の規格化のため。 $p > 0.5$ のとき、ON の割合が高くなる。(c) 疎通バイアス: ある入力変数を 0 か 1 に定めただけで、関数の出力が他の入力変数に依らずに決まってしまう時、その入力変数を疎通入力と呼ぶ。関数 A では、 $x_j = 1$ でありさえすれば、 x_k に依らずに $f_i = 0$ であるので、 x_j は関数 A の疎通入力である。一般に、 K 個の入力変数を持つ関数は、0 個から K 個の疎通入力を持ち得る。この疎通入力数の割合を変化させるバイアスを疎通バイアスと呼ぶ。 p バイアスとは違い、1 つのパラメータだけで決定することはできないので、他にも取り決めが必要である。

軸は取り得るブール関数に対応している。確率分布 $P_i(f_i)$ の形を一様分布から変化させることで、ブール関数の選び方にバイアスをかけることができる。図 3(b) と (c) のバイアスは、カオス-秩序転移を引き起こしうることが知られている [3]。

2.2 固定点の数

ネットワークトポロジーとすべてのノードのブール関数 $\{f_i; i = 1, \dots, N\}$ が一意に与えられた系を考える。固定点においては、すべてのノード i で $x_i(t)$ と $x_i(t+1)$ が等しいので、系に含まれる固定点の数は次の式で定まる:

$$n(\{f_i\}) = \sum_{x_1=0,1} \dots \sum_{x_N=0,1} \prod_{i=1}^N \delta[x_i, f_i(\{x_j; j \in J_i\})].$$

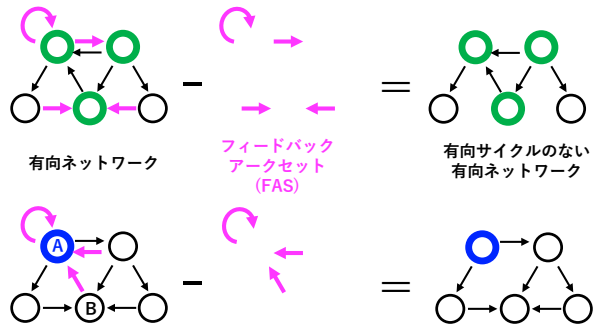


図 4: (左) 有向ネットワーク。上と下は同じものである。(中) FAS。(右) FAS を取り除いたもの。FAS の定義により、有向サイクルは消える。ここで、図内の下の例を用いて、定理に当てはまる状況を考えてみる。ノード A の確率分布 $P_A(f_A)$ を FAS が確率中立になるよう (最も簡単には一様分布) に取りさえすれば、(ノード B の 3 変数関数含め、他のノードの関数を様々に選べるにも関わらず) 直ちに $\langle n \rangle = 1$ となる。

ここで、 $\delta[a, b]$ は、クロネッカーのデルタであり、 $a = b$ のとき $\delta[a, b] = 1$ 、 $a \neq b$ のとき $\delta[a, b] = 0$ を出力する。

与えられたネットワークトポロジーの下で、各 f_i が $P_i(f_i)$ に従って分布するときの、固定点の数の期待値は、

$$\langle n \rangle = \sum_{f_1} \dots \sum_{f_N} n(\{f_i\}) \prod_{i=1}^N P_i(f_i), \quad (2)$$

と表される。 \sum_{f_i} は、与えられたネットワークにおいて f_i が取り得るすべてのブール関数についての和を表す。式 (2) を見ると、 $\langle n \rangle$ は、ネットワークトポロジーとブール関数の確率分布の両方に依存すると予想される。この予想は次の定理で覆される。

3 定理

定理— 確率的に中立なリンクの集合が Feedback Arc Set (FAS) であるならば、 $\langle n \rangle = 1$ である。

3.1 用いられた概念

ここでは定理の証明に用いた概念の説明をする²。これらに加えて、記号の用法さえ押さえれば、計算の理解は容易なので、証明の詳細は文献 [4] を見ていただきたい。

N-同値関係: x_j を入力変数の 1 つとして持つ、ある関数 $f_i(x_j, \{x_k; k \in J_i - j\})$ を考える。 j は、 $j = i$

² 「確率中立」は文献 [4] で新しく導入した。「N-同値関係」と「FAS」は従来からある。

でも $j \neq i$ でもよい。 f_i と同じ入力変数を持つ別の関数 g_i が、 $g_i(x_j, \{x_k; k \in J_i - j\}) \equiv f_i(x_j, \{x_k; k \in J_i - j\})$ と書かれるとき、 f_i と g_i は j に関して N-同値関係³にあるという [5]。図 3 の関数 A と C は、入力変数 x_j の 0 を 1 に、1 を 0 に入れ替えた関係になっており、 j に関して N-同値である。関数 A と B は、 k に関して N-同値である。

なお、 f_i が x_j の単調増加関数のとき、 j に関して N-同値な関数である g_i は、 x_j の単調減少関数となる。ここで、 x_j の単調増加関数とは、(図 3 の関数 C のように) x_j 以外の入力変数の値を固定して $x_j: 0 \rightarrow 1$ としたとき、出力が増えるか変わらないかのどちらかを必ず満たす関数のことである。同様に、単調減少関数とは、出力が減るか変わらないかのどちらかを必ず満たす関数である。さらに、 f_i が x_j の単調増加関数のとき、リンク $j \rightarrow i$ は活性の役割を果たしていると解釈でき、正の符号を持つと定義される。単調減少関数のときは、抑制の役割で、負の符号を持つ。

確率中立: ノード j から入力を受けるノード i の確率分布 P_i が、 j に関して N-同値な関数のペア (f_i, g_i) すべてに対して、 $P_i(f_i) = P_i(g_i)$ を満たすとき、リンク $j \rightarrow i$ は確率的に中立であると呼ぶ。リンク $j \rightarrow i$ が確率的に中立であるとき、そのリンクが正である確率と負である確率は等しい。ただし、確率中立の概念は、リンクの正負が決まらない非単調関数のペア (例えば、図 3 の関数 I と J) に対しても適用できる。図 3(a-c) で示されている種類の確率分布を考えている限り、すべてのリンクは確率中立である。

FAS: 有向ネットワークにおいて、すべての有向サイクルの少なくとも 1 本ずつのリンクを含む任意のリンクの集合として定義される [6]。つまり、有向ネットワークから FAS を取り除くと、有向サイクルは消失する (図 4)。一般に、FAS は唯一でない。

3.2 定理の意味

固定点の数の期待値は、 2^N の広さをもつ状態空間の中で、定理の条件を満たしている限りわずかに 1 である。 n は非負であるので、例えば $n = 100$ が実現するような確率は 1% 以下であると断定される。この意味で、たいていの場合、固定点の数のオーダーは 1 である。図 3(b) と (c) のようなカオス-秩序転移を引き起こすバイアスのもとでは、すべてのリン

クが確率中立であることから、ネットワークポロジをどのように調節しても、固定点数の期待値は何ら変わらない。定理は、部分的な情報だけから期待値を決定しており、例えば図 4 では、ノード A の確率分布 $P_A(f_A)$ が適当な条件 (例えば一様分布) を満たしているだけで、他のノードがどのような確率分布を取ろうとも、 $\langle n \rangle = 1$ が言える。

3.3 確率中立を破る

定理より、 $\langle n \rangle \neq 1$ となるための必要条件は、サイクル上のすべてのリンクで確率中立が破れているようなサイクルが 1 つ以上存在することである。そこで、互いに交わらない ν 個のサイクル上のリンクすべてで、確率中立を対称に破っているような状況 (図 5) を考える。ただし、 ν 個のサイクルは、サイクル内にショートカットを持たないものとする。それ以外のリンクについては、確率中立を仮定する。このとき、 $\langle n \rangle$ は

$$\langle n \rangle = \prod_{\eta=1}^{\nu} (1 + \text{「サイクル } \eta \text{ の効果」}), \quad (3)$$

のような形で表現できる (詳しくは文献 [4] を参照)。「サイクル η の効果」は、 -1 から 1 までの値を取る。確率中立を対称に破るために導入するパラメータを分かりやすく取ると、「サイクル η の効果」の符号は容易に知ることができる。符号が正のときは、そのサイクル内では、ポジティブフィードバックがネガティブフィードバックよりも確率的に優勢であることを意味し、 $\langle n \rangle$ を増やす。負のときは、その逆である。BN におけるポジティブ (ネガティブ) フィードバックとは、有向サイクルを構成するリンクの符号

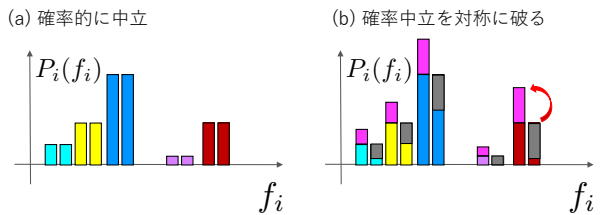


図 5: 確率中立と確率中立を対称に破った場合の確率分布の概念図。(a) リンク $j \rightarrow i$ が確率的に中立だとすると、 j に関する N-同値な関数のペアは等確率である。(b) すべてのペアで、一定量 (ピンクの部分) の確率を、 x_j の単調減少関数から単調増加関数へ移動させると、リンク $j \rightarrow i$ の確率中立が対称に破れ、リンクが正である確率が高くなる。

³Negation (否定) 同値関係

の積が正(負)になることである。ポジティブフィードバックの数が、 n そのもの上限を与えること、つまり、大きな n のための必要条件であることは以前から知られていた [7]。ここでは、 $\langle n \rangle$ を増やす十分条件を示したことになる。

4 まとめ

本稿では、任意のネットワークポロジーマトリクスをもつ BN に対して、「FAS が確率的に中立であれば固定点数の期待値は 1 である」という定理を解説した。概して言えば、固定点はたいていの場合で少ない。固定点を増やすためには、ネットワークの調整や各ノードに対するバイアスの導入を独立に行っているだけでは効果がなく、うまくポジティブフィードバックがかかるような系を構築しなければならない。

証明で用いられた計算のアイデアは、他の局面にも活用できる可能性がある。2 状態を取る要素から成るネットワーク系を改めて BN と捉え直すことで、他の期待値の計算やそのネットワーク依存・非依存性の理解が進展することにも期待したい。

参考文献

- [1] A. Mochizuki, *J. Theor. Biol.* **236**, 291 (2005).
- [2] P. M. de Espanés, A. Osses, and I. Rapaport, *Biosystems* **150**, 167 (2016).
- [3] S. A. Kauffman, *Investigations* (Oxford University Press, 2000).
- [4] F. Mori and A. Mochizuki, *Phys. Rev. Lett.* **119**, 028301 (2017).
- [5] T. Sasao, *Switching theory for logic synthesis* (Springer Science & Business Media, 2012).
- [6] J. Bang-Jensen and G. Z. Gutin, *Digraphs: theory, algorithms and applications* (Springer Science & Business Media, 2008).
- [7] J. Aracena, *B. Math. Biol.* **70**, 1398 (2008).