数理モデル化と問題解決 42－19
（２００２．１１．２９）

セル・オートマタによる符号化手法とその分析

福原　義久*　　　武藤　佳恭**

概要

セル・オートマタ(以下 CA)とは有限オートマトンを空間内に並べて配置し，近接するセル同士を結合したシステムである．CA は並列・同期的に動作し，空間内の状態はオートマトンの動作規則によってさまざまに変化する．通常，CA は一つの系で状態変化が完結しているが，本研究では複数の CA システムを同期して動作させることにより，各 CA 空間の情報量を下げ，符号化の手法として用いることができることを示した．また，どのような遷移規則の場合に効果的に符号化が行えるかを遷移規則の持つパターンの状態をあらわすパラメータを定義して，詳細に分析した．

Information encoding by using cellular automata

Yoshihisa Fukuhara* and Yoshiyasu Takefuji**

Abstract

In this paper, we propose a new crossover operator for the genetic algorithms by using a deterministic crossover length. We confirm the search ability of the proposed method by tree benchmark problems and two kinds of NP-complete problems : Ramsey problem and TSP. By define a crossover length in advance, a search ability of GA has significant increased. Especially, the proposed method is effective in 2-point crossover.

1．はじめに

セル・オートマタ(以下 CA)[1]は生命現象，化学現象，社会現象などの振る舞いを考える上でのもっともシンプルかつ汎用的なモデルである．現実世界のアナロジーとしてCAを用いてさまざまな種類の研究が行われている一方で，CA 自体の性質や振る舞いについての研究も盛んに行われてきた．CA を情報処理プロセッサとしてとらえたとき，二次元のCAが万能チューリングマシンとして動作することが Berlekamp[2]らによって示された．その後，Wolfram[3]は１次元 CA の動作を詳細に検証した結果，秩序的な状態からカオス的な状態に至る 4 つの状態に分けられることを示した．彼はこのような複雑な状態を示すCAには計算万能性があるのではないかと考えた．このように，シミュレータとしての用途や数理的分析の研究が盛んに行われている一方で，CA を用いて具体的な問題解決を行わせる例はそれほど多く見当たらない．これは CA がその単純な動作原理とはうらはらに予測困難で複雑な振る舞いを示すことに起因しているのだろう．

CA の示す複雑な振る舞いは，計算万能性や創発現象，生命現象との関係が述べられるほどの可能性[4]を持ちつつも，具体的に利用することが困難なのである．我々は，なるべく CA の本質を失わせることなくその情報処理の可能性を検証したい．そこで本研究ではCAを用いた符号化装置を考案した．提案システムは，複数の CA を同時に実行し，各時間でそれぞれの状態を比較しながら情報の次元数を低下させていくことによって成される．提案システムは，CA の本質的な特徴を損なうことなく符号化という最も基本的な情報処理を行うことに成功したものである．また，提案システムがどのような遷移規則のときに効果的に動作するのかを遷移規則のパターンに対する応答特性を示す新たなパラメータを定義し分析を行った．

2．手法

2.1 一次元セル・オートマトン

CA は離散的時空間の中で次の時刻でのあるセルの状態が，そのセルと近傍のセルの状態によって決定される．最も基本的な 2 状態 3 近傍の一次元 CA の場合，自セルと両隣の 3 つのセルの状態から次の時刻での自セルの状態が決定される．2 状態 3 近傍の場合，図 1(1)に示されたような 8 つのエントリーの組み合わせからなる遷移規則が必要であ

*慶應義塾大学 政策・メディア研究科 武藤 佳恭研究室 Takefuji Laboratory, Graduate School of Media and Governance, Keio University, Japan
**慶應義塾大学　環境情報学部　Faculty of Environmental Information, Keio University, Japan

る．遷移規則の組み合わせの数は k を状態数，n を自セルを含めた近傍の数とすると k^{k^n} 個存在する．一般に遷移規則を f であらわすとすると，セルの状態 a は式 1, 2 であらわされる．ここで $a_{i,j}^t$ は時刻 t における系 i の j 番目のセルの状態をしめしている．一次元の任意の初期パターンにこの遷移規則を適用することによって，パターンはさまざまに変化していく．

$$a_{i,j}^{t+1} = f(a_{i,j-x}^t, \ldots, a_{i,j}^t, \ldots, a_{i,j+x}^t) \tag{1}$$

$$x = (n-1)/2 \tag{2}$$

2.2 CA を用いた符号化装置

本研究での符号化とは，複数の異なる有限の数列をあるアルゴリズムにしたがって，より低次元のそれぞれ異なる数列に写像することを意味する．

まず符号化を行いたい複数のパターンに対してそれぞれ遷移規則を適用する．提案システムの動作は，これら各 CA 系からの出力を各時間でそれぞれ比較し，不要なセルを消滅することによって情報の次元を下げていくことによってなされる．つまりプロセスの進行により空間からセルを取り除いてやる必要があるのだが，これはいいかえれば，セルのとる値のほかにそのセルが使用可能かどうかのフラグを追加することと同じである．すべての系を通して，各セルが使用可能かどうかは式 3 によって導かれる．ここで S_j^t は m 個の CA 系が存在するときの t 時における j 番目が使用可能かどうかのフラグを示す．S が真の a の集合に対してのみ式 1 が適用される．本研究では関数 f を a がすべて同じ値のとき使用不可能，それ以外のとき使用可能な状態を返すと定義した(式 4)．座標 j のセルの生き死にはすべての CA 系において共通である．つまりある系でのセル j は使用可能にもかかわらず，別の系でのセル j が使用不可能であるということはありえない．また，それぞれの CA 系は同じ遷移規則で動作するものとする．

$$S_j^t = f(a_{1,j}^t, a_{2,j}^t, \ldots a_{i,j}^t) \tag{3}$$

$$f(x) = \begin{cases} 0 & a_{1,j}^t = a_{2,j}^t \ldots = a_{i,j}^t \text{のとき} \\ 1 & \text{それ以外のとき} \end{cases} \tag{4}$$

上記の符号化プロセスについて具体例を用いて説明する．図 1 の模式図の例では，入力パターンとして 2 状態 8 ビットのパターンを 2 つ用意した．2 状態のパターンを 2 つ分類する作業なので，最低 1 ビットの情報量が必要であるから目標とする次元数は 1 とする．

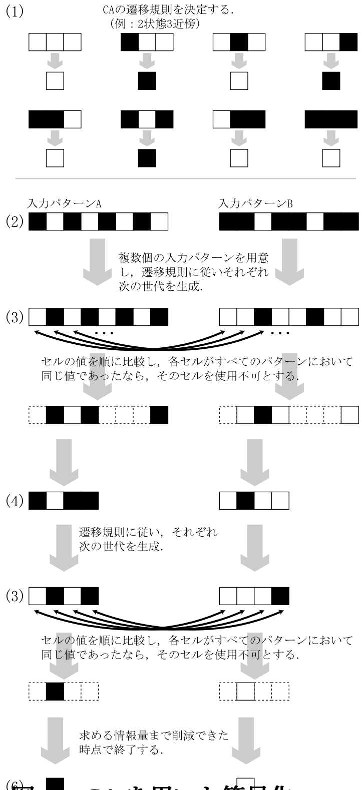

図 1　CA を用いた符号化

1. CA の遷移規則を適当に決める．図の例では 2 状態 3 近傍の遷移規則を決定した．
2. 長さの同じ複数の入力パターンを用意し，それぞれから遷移規則に従って次世代のセルを生成する．
3. 生成されたパターンのセルを同じ座標ごとにそれぞれ比較し，すべて同じ値であった場合にはそのセルを使用不可とする．例では，1，5，6，7 番目のセルが 2 つのパターンで同じ値であるため使用不可能とする．
4. 残ったセルに再度遷移規則を適用する．ただしここで用いる近傍のセルとは**使用可能なセルのうち最も近くにあるもの**を意味する．図中では説明のため，残ったセルでパターンを再度構成した．
5. (3)から(4)を期待する次元にセルの数が減るまで繰り返す．
6. 期待する次元数まで情報量が下がった時点で終了する．

しかし，以上の操作は学習などを伴わず，あくまで遷移規則だけに基づいた決定論的操作なので

必ずしも望む結果が得られるとは限らない．入力パターンを図の例のように正しく符号化できるか否かは入力パターンと遷移規則との複雑な関係により事前に予測することは困難である．最終的にシステムのとりうる状態は以下の 4 つが考えられる．

1. 期待する次元数までセルの数が減り符号化が成功する．
2. 期待する次元数までセルの数が減るが，正しく符号化できていない．
3. 期待する次元数よりもセルの数が減ってしまう．もしくはセルがすべて使用不可になる．
4. 期待する次元数に達する前にリミットサイクルに陥り，それ以上セルの数が減らなくなる．

提案手法では最終的な出力が望まれるものであった場合，その遷移規則と各時刻での使用可能なセルの座標を記録してさえおけば適用したパターンは決定論的操作により必ず分類することができる．つまり単純かつ高速なパターン分類器として用いることもできるのである．また提案手法は二次元以上の CA にも適用可能である．しかし実用的なアプリケーションとして提案手法にアプローチする前に，まずどのような状態の CA が符号化能力をもちうるのかについて検証する必要があるだろう．次章以降では，さまざまな遷移規則と入力パターンを用いて提案手法を検証し，遷移規則と符号化の能力の関係について考察する．

3. 評価手法

我々は前述のようなシステムがどのような遷移規則の場合効率よく動作するかを知りたい．ここではまず Langton[4]の提案したλパラメータ(式 5)

$$\lambda = (k^n - n_q)/k^n \tag{5}$$

を用いる．

kは状態数であり，nは状態遷移にかかわる近傍セルの数である．n_q は静状態を示す．静状態とは，セルがとりうるいくつかの状態のうち任意に定められたある一つの状態のことである．しかしλの解像度は k^n に等しく，CA の持つ全遷移規則数 k^{k^n} に対して粗いといわざるをえない．そこで本論文では遷移規則のパターンに対する応答の振る舞いを検証するために新たにパラメータ G を定義した．Gは式 6，7 および 2 によって定義される．

ここで，pは遷移規則一組の中でのエントリーを

$$G = \sum_{p=1}^{k^n} \sum_{q=-x}^{x} qCpqC'_p \tag{6}$$

$$C = \begin{cases} 0 & \text{セルが静状態のとき} \\ 1 & \text{セルが静状態以外のとき} \end{cases} \tag{7}$$

ナンバリングしたものであり，qは各エントリーにおいての自セルを原点とした各セルの座標である．C_{pq} はp番目のエントリーでの各セルの状態を示し，C'_p はp番目のエントリーから生成されるセルの状態を示す．また，近傍数 n は奇数であることを前提とする．

G の値は，CA の状態・近傍数に応じて固有の負から正の値を持ち，0 のとき CA がエントリーのパターンに対して対称的な応答をすることを示す．Gは喩えていうならば遷移規則のパターンに対する応答の重心といえるだろう．これらλと G の二つの観点から評価することにより，CA の遷移規則と符号化能力の関係についてより詳細に分析することができる．

4. 実験結果

2 状態 5 近傍の CA を用いて，遷移規則と符号化能力についての検証を行った．実験は一つの遷移規則につき，ランダムに生成された 100 ビットの入力パターンのセットを 1000 回与え，符号化成功回数を測定した．遷移規則は各λの値につき 200 個をランダムに抽出したものを用いた．

図 2～5 は，各λ，Gにおける符号化能力を示したものである．

5. 考察

分析の結果，提案手法では以下の 3 つの特徴を持つ遷移規則の符号化能力が高いという傾向を得た．

- 遷移規則のエントリーのパターンに対する応答が非対称な場合．
- λパラメータが約 0.2 から 0.3.
- 遷移規則がクラス II に属する．

ただしλや G が結果にたいして相関を持っていると思われるものの，状態・近傍数および入力パターン数に応じてその影響の度合いが変化する場合がある．また，パラメータで指定した遷移規則のうちのごく一部だけが符号化能力を有しており，すべてのケースにおいて普遍的かつ的確に符号化能力の高い遷移規則を抽出する基準は発見できなかった．一方，以下の二点の特徴のうちどちらかを有している場合はすべての実験に共通して符号化能力が低いことが判明した．

- λ=1-1/kでなおかつ G=0 に近いとき．

● λないし G がそれぞれがとりうる最大・最小値に近いとき.

以上のことから，符号化能力には入力に対する応答の非対称性ならびに遷移規則のパターンに対する非対称的応答性が関与していると推測される．ところで Wolfram はクラス IV において，計算万能性についての推測をたてた[3]．同様にカオスの縁と計算能力の関係について指摘するものもある[5]．しかし今回の実験結果からは，λパラメータからみるとクラス IV に近い場合もあるが，実際のパターンはクラス II や III に属するものであり，関連性は確認できていない．

6. 結論

本研究では，CA を用いた符号化手法を提案した．提案手法は決定論的操作で簡便かつ高速に符号化を行えるものである．提案手法は二次元以上および二値以上のパターンに対しても適用可能である．符号化可能な遷移規則を用いれば，簡単なパターン分類器として用いることが可能である．提案手法を具体的なアプリケーションに応用するには，改善しなければならない問題や検証しなければならない問題が多いが，CA を用いた情報処理に一つの可能性を示すものとして提示したい．特に提案手法は入力と CA の遷移規則の関係からすべての場合で正しく符号化できるとは限らないという制限がある．本研究では Langton の提案したλパラメータと新たに定義したパラメータ G を用いてこれを検証し，入力に対する応答の非対称性ならびに遷移規則のパターンに対する非対称的応答性が符号化能力に寄与しているのではないかとの推測をおこなった．

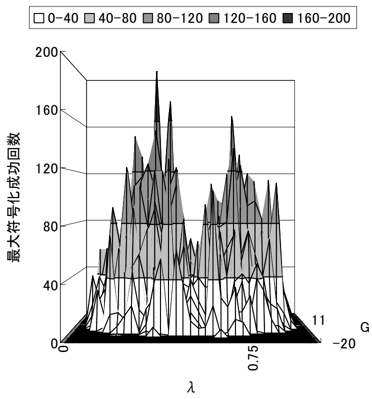

図 2 λからみた最大符号化回数

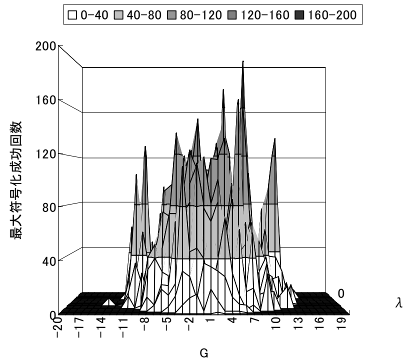

図 3 G からみた最大符号化回数

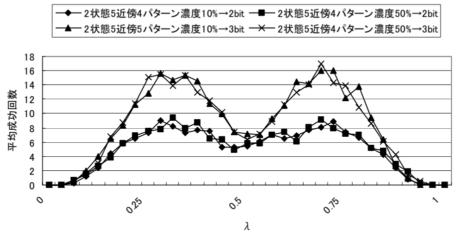

図 4 λからみた平均符号化回数

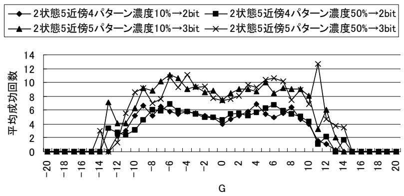

図 5 G からみた平均符号化回数

References

[1] von Neumann, J. : Theory of self-reproducing automata, edited and completed by A. Burks, University of Illinois Press (1966).

[2] Berlekamp, E., Conway, J. and Guy, R. : Winning Ways for your Mathematical Plays, Academic Press (1982).

[3] Wolfram, S.: Universality and Complexity in Cellular Automata, Physica D, 10, pp. 1--35 (1984).

[4] C.G.Langton: Computation at the Edge of Chaos: Phase Transitions and Emergent Computation, Physica D, 42, pp. 12--37 (1990).

[5] Packard, N. H.: Adaptation toward the edge of chaos, Dynamic Patterns in Complex Systems, Singapore, World Scientific, pp. 293--301 (1988).