Análise Estatística de Geradores Pseudo-Aleatórios Baseados em Autômatos Celulares

Nicoli Pinheiro de Araújo, Elloá B. Guedes

[1]Núcleo de Computação
Escola Superior de Tecnologia
Universidade do Estado do Amazonas
Av. Darcy Vargas, 1200 – Manaus – Amazonas

npda.eng@uea.edu.br, ebgcosta@uea.edu.br

Resumo. *Autômatos celulares são modelos de computação baseados em células que se auto-reproduzem. De acordo com a literatura, estes autômatos são estruturalmente simples, mas capazes de gerar padrões complexos, o que culmina na adoção dos mesmos para geração de números pseudo-aleatórios. Levando isto em consideração, este trabalho teve como objetivo analisar a qualidade estatística das sequências produzidas por geradores baseados nas versões elementares e totalísticas destes autômatos. Como resultado, foi possível constatar que este modelo de computação não se mostra adequado para geração de sequências pseudo-aleatórias, mostrando baixa qualidade estatística. Estes resultados impactam na não-indicação deste modelo para geração de sequências numéricas, pois podem comprometer as aplicações que os utilizam.*

Abstract. *Cellular automata are computational models based on cells that self-reproduce. According to the literature, these automata are simple-structured yet able to yield complex patterns, which implies in their adoption as pseudorandom generators. Taking that into account, the main objective of this work was to analise the statistical quality of pseudorandom sequences produced by the elementary and totalistic version of such automata. As a result, it was possible to verify that this model of computation is not adequate for pseudorandom sequence generation, as it shows a low statistical quality. These results indicates that the adoption of such pseudorandom generators is not recommend because they can harm the applications in which they are used.*

1. Introdução

Autômatos celulares são um modelo computacional caracterizados por um vetor n-dimensional de células que se auto-reproduzem. Eles exemplificam componentes matemáticos estruturalmente simples, mas capazes de gerar padrões complexos [Wolfram 2002]. Este tipo de autômato tem sido utilizado em diversas aplicações, tais como na modelagem de sistemas biológicos, reações químicas, fractais, comportamento de fluidos, dentre outros [Schiff 2010, Salcido 2013].

Uma das aplicações dos autômatos celulares, em particular, é a geração de sequências pseudo-aleatórias. Em seu trabalho seminal, Wolfram descreveu maneiras de relacionar os vetores de bits que compõem os autômatos celulares com números, permitindo a geração de sequências numéricas [Wolfram 1986, Gentle 2003]. A partir deste trabalho, diferentes estratégias para produzir sequências pseudo-aleatórias a partir de autômatos celulares foram propostas na literatura [Kang et al. 2008, Bardell 1990, Guan and Tan 2004].

Porém, antes de usar sequências pseudo-aleatórias, é essencial garantir a qualidade das mesmas, isto é, estabelecer uma medida de quão próximas as propriedades das saídas de um gerador pseudo-aleatórios são próximas de um gerador aleatório ideal, baseado em um processo uniforme independente e identicamente distribuído [Gentle 2003]. Vários testes estatísticos endereçam essa questão, pois a aleatoriedade é uma propriedade probabilística, ou seja, as propriedades de um gerador ideal podem ser descritas em termos de probabilidades. Isto significa que é possível descrever *a priori* as saídas de um gerador ideal. Existem diversos testes estatísticos capazes de verificar a presença ou ausência desses "padrões", os quais podem indicar se uma sequência numérica oriunda de um gerador ideal ou não [L'Ecuyer 1992, Rukhin et al. 2008].

Considerando a importância de testes estatísticos neste cenário, o número de referências encontradas na literatura sobre a qualidade estatística de geradores pseudo-aleatórios baseado em autômatos celulares ainda é incipiente, especialmente se considerada a proposição original feita por Wolfram [Wolfram 1986]. Com o intuito de melhorar o conhecimento sobre estes geradores, o trabalho em questão mostra os resultados e conclusões de um conjunto de testes estatísticos aplicados a diversos geradores pseudo-aleatórios baseados em autômatos celulares elementares e totalísticos. Os resultados obtidos mostram evidências de que, apesar do comportamento complexo resultante de regras simples dos autômatos celulares, a utilização desse modelo de computação para geração de sequências pseudo-aleatórias resulta em falha na maioria dos testes estatísticos submetidos.

Para apresentar e discutir os resultados obtidos, este trabalho está organizado como se segue. Uma breve apresentação de autômatos celulares enquanto geradores pseudo-aleatórios é mostrada na Seção 2. As estratégias para avaliar a qualidade de geradores pseudo-aleatórios é mostrada na Seção 3. Os resultados destas avaliações aplicadas aos geradores pseudo-aleatórios é mostrada e discutida na Seção 4. Por fim, considerações finais e trabalhos futuros são apresentados na Seção 5.

2. Autômatos Celulares

Autômatos celulares são um modelo computacional que representam muitos sistemas naturais. A versão unidimensional destes autômatos é composta de um vetor de células idênticas, discretas e que podem assumir valores de um conjunto finito. O valor de cada célula é modificado em passos discretos do tempo de acordo com uma regra determinística, a qual faz uso do próprio estado da célula e do estado das células vizinhas para determinar o novo estado de uma célula, caracterizando a evolução de um autômato celular [Schiff 2010]. A Definição 1 formaliza este modelo de computação.

Definição 1 (**Autômatos Celulares [Wolfram 1984]**). Um autômato celular unidimensional consiste em uma linha de células, e é definido por três parâmetros (r, k, ϕ). O valor de a_i referente à célula i é atualizado em passos discretos do tempo de acordo com a seguinte expressão:

$$a_i^{(t+1)} = \phi \left[a_{i-r}^{(t)}, a_{i-r+1}^{(t)}, \ldots, a_{i+r}^{(t)} \right], \tag{1}$$

em que o parâmetro k informa os possíveis valores que as células podem possuir, especificados no intervalo $[0, k[$; r indica a quantidade de células vizinhas que são utilizadas na determinação do novo valor de uma célula; e, ϕ, por sua vez, é uma regra determinística.

Existem diversas variantes de autômatos celulares unidimensionais, dentre as quais se destacam os elementares e os totalísticos. No caso dos elementares, as regras de coloração das células baseiam-se em códigos construídos a partir de números inteiros, enquanto no caso dos totalísticos a coloração baseia-se em uma média das cores das células vizinhas [Wolfram 2002].

Considerando a definição de autômatos celulares e suas duas variantes apresentadas, Wolfram [Wolfram 1984] identificou quatro tipos de padrões produzidos na saída destes autômatos. São eles:

1. Padrões que desaparecem com o tempo;
2. Padrões que evoluem para um certo tamanho fixo;
3. Padrões que crescem indefinidamente em uma velocidade fixa;
4. Padrões que crescem e se contraem de maneira irregular.

A Figura 1 ilustra exemplos de autômatos celulares que se classificam nos tipos de padrões apresentados. O autômato celular da Figura 1a é do tipo elementar com regra numérica igual a 0. Este autômato é inicializado com um único site na cor preta, o qual desaparece após a segunda iteração. O autômato celular da Figura 1b também é do tipo elementar e possui regra numérica igual a 12, repetindo o padrão inicial ao longo do tempo. O autômato da Figura 1c é o *Rule 30*, cujo padrão cresce ao longo do tempo. Por fim, o autômato celular da Figura 1d, de regra totalística 600 e $k = 3$, mostra um padrão que cresce e se contrai.

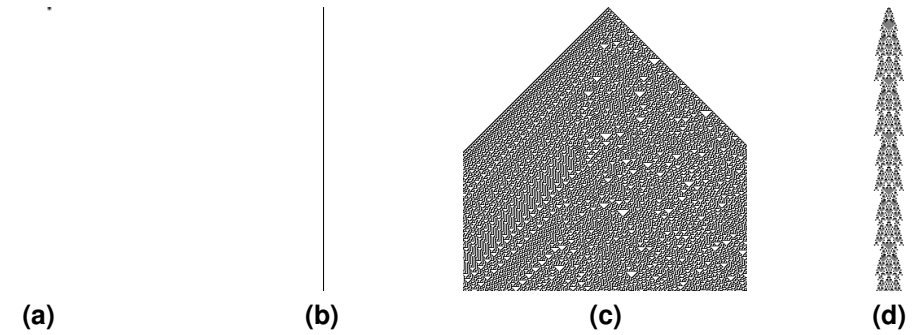

Figura 1: Ilustração dos diferentes padrões gerados por autômatos celulares.

Com respeito aos autômatos celulares elementares, em particular, existem 256 autômatos deste tipo, dos quais 65% produzem padrões que permanecem em um tamanho fixo (Tipo-2). Apenas 14% deles exibem padrões mais complexos (Tipo-3). Nenhum dos autômatos celulares elementares é capaz de gerar padrões que crescem e se contraem. Este tipo de padrão só pode ser observado em certos tipos de autômatos celulares totalísticos [Wolfram 2002].

Considerando os autômatos celulares como geradores pseudo-aleatórios, a descrição do valor de um site mostrada na Eq. (1) da Definição 1 mostra como produzir bits a partir de autômato celulares [Wolfram 1986]. A inicialização do autômato celular funciona como a semente de um gerador pseudo-aleatório; a regra ϕ é o processo iterativo que produz os números seguintes; e o valor das células são a saída do gerador [Gentle 2003]. Assim, de acordo com esta caracterização, cada regra de uma variante de um autômato celular corresponde a um gerador pseudo-aleatório.

Esta caracterização de gerador pseudo-aleatório a partir de um autômato celular é oriunda do trabalho seminal de Wolfram [Wolfram 1986]. Entretanto, trabalhos recentes da literatura mostram outras estratégias para gerar sequências pseudo-aleatórias a partir de autômatos celulares [Kang et al. 2008, Bardell 1990, Guan and Tan 2004]. Neste trabalho será considerada a abordagem de Wolfram para geração de sequências pseudo-aleatórias a partir de autômatos celulares, pois é a base utilizada em muitas adaptações requeridas pelos trabalhos mais recentes.

Em termos de aplicações, autômatos celulares são amplamente utilizados em modelagens que envolvem auto-reprodução, padrões complicados e certa aleatoriedade. Por exemplo, são utilizados na modelagem do crescimento competitivo de espécies de algas [Chen et al. 2002]; na detecção de intrusos em redes sem fio [Navid and Aghababa , Cap. 5]; na remoção de ruídos e detecção de bordas em imagens digitais [Popovici and Popovici 2002]; no reconhecimento de padrões [Navid and Aghababa , Cap. 3]; na modelagem da evolução de certas vegetações [Ye et al. 2010], dentre outros.

3. Qualidade Estatística de Geradores Pseudo-Aleatórios

Muitos métodos estatísticos utilizados pela Engenharia e Ciências Naturais demandam utilização de *números aleatórios* para simulação de processos físicos e biológicos. A utilização destes números também se aplica à outros domínios, por exemplo, a criptografia – para permitir a troca segura de dados, e aos sistemas operacionais – possibilitando a execução de determinados algoritmos.

Para obtenção de números aleatórios normalmente se utiliza uma fonte geradora de números aleatórios, que consiste em um *hardware* especializado para a captura de números aleatórios gerados em decorrência de algum fenômeno físico, por exemplo, o decaimento de uma substância química. Apesar dos números resultantes serem verdadeiramente aleatórios, o custo deste hardware e a imprevisibilidade dos fenômenos físicos costumam ter um impacto negativo na adoção destes geradores.

Em virtude dos fatores mencionados, a utilização de números aleatórios se torna inviável para algumas situações. Porém, algumas delas, sem prejuízo aos resultados, podem recorrer à utilização de *números pseudo-aleatórios*: sequências de números obtidas a partir de funções matemáticas, executadas por computadores convencionais, que assemelham-se à sequências de números aleatórios de uma dada distribuição de probabilidade [Gentle 2003].

A geração de números pseudo-aleatórios possui várias vantagens em relação à geração de números aleatórios, pois (i) costuma ser mais rápida; (ii) não depende de hardware específico, (iii) pode ser facilmente configurada; (iv) várias técnicas diferentes para geração deste tipo de número já foram desenvolvidas e (v) está presente na maioria das linguagens de programação e bibliotecas de sistemas operacionais.

Apesar do grande conjunto de vantagens práticas, quando certas sequências de números fornecidas por geradores pseudo-aleatórios são analisadas por meio de testes estatísticos, tais sequências podem fornecer indícios de que seus respectivos geradores podem ser inadequados para determinadas aplicações, por exemplo, na criptografia, em que a previsibilidade da sequência pode representar uma quebra na segurança dos dados criptografados [Williams and Clearwater 1997].

Assim, a escolha do gerador adequado para uma determinada aplicação não depende somente dos atributos deste gerador (técnica para geração, intervalo dos números geradores, período, etc.), mas também da qualidade dos atributos das sequências produ-

zidas por estes geradores. Alguns destes atributos são a frequência dos bits 0's e 1's, a ausência de padrões sistemáticos, a correlação nula entre as amostras produzidas, etc. Uma vez que todos os atributos analisados são relativos às sequências produzidas pelos geradores, é necessária uma *metodologia empírica* de verificação de tais atributos nas sequências produzidas por um gerador.

De acordo com esta metodologia empírica, inicia-se com um teste de hipóteses, cuja hipótese nula afirma que a sequência numérica que está sendo testada é aleatória. A hipótese alternativa enuncia o oposto da hipótese nula. Por meio de testes estatísticos, é possível derivar uma conclusão que culmina em aceita ou rejeitar a hipótese nula. Para tanto, o usuário deve primeiramente determinar o nível de confiança nos resultados, denominado α. Tipicamente os valores de α encontram-se no intervalo $[0.001, 0.01]$. Em seguida, o teste é executado sobre uma determinada sequência, produzindo uma estatística, chamada p-valor e que é uma função dos dados de entrada fornecidos. Se o p-valor for menor que α, diz-se que a hipótese nula foi rejeitada. Em caso contrário, diz-se que a hipótese nula não foi rejeitada. Esta conclusão sobre a sequência impacta no gerador que a produziu, refletindo a qualidade do mesmo [Rukhin et al. 2008].

Embora os testes gerem conclusões a partir da sequência fornecida como entrada, não constituem uma prova formal que ateste aleatoriedade na geração da sequência nem tampouco resultam em uma conclusão definitiva, mas sim probabilística [Menezes et al. 1996]. Levando isto em consideração, quando testes estatísticos são realizados, três hipóteses são consideradas sobre as sequências binárias que lhes são fornecidas como entrada:

1. **Uniformidade**. Em qualquer ponto da geração da sequência, a ocorrência de 0's e 1's é igualmente provável, ou seja, a probabilidade esperada de cada um é exatamente $1/2$;
2. **Escalabilidade**. Qualquer teste aplicado à sequência pode ser aplicado à subsequências escolhidas aleatoriamente. Se a sequence é aleatória, as subsequências também devem sê-lo. Assim, espera-se que qualquer subsequência também passe em qualquer teste estatístico quando a sequência original também passa;
3. **Consistência**. O comportamento do gerador deve ser consistente mesmo considerando diversas inicializações (sementes). É inadequado testar um gerador pseudo-aleatório considerando resultados de uma única inicialização.

Os testes estatísticos propostos para avaliar geradores pseudo-aleatório são diversos, conforme documentado na literatura [Menezes et al. 1996, Gentle 2003, Rukhin et al. 2008, L'Ecuyer 1992, Knuth 1998], e não há um consenso sobre um conjunto de testes efetivos para as hipóteses previamente mencionadas. Em termos práticos, utilizam-se diversas baterias de testes disponíveis em softwares como o NIST [Rukhin et al. 2008], Diehard [Marsaglia], Sieve [Guedes et al. 2009], dentre outros.

No escopo deste trabalho será tomado como referência o conjunto de testes estatísticos disponível na ferramenta open-source e gratuita Sieve [Guedes et al. 2009]. O Sieve é capaz de testar sequências produzidas por geradores de números aleatórios e pseudo-aleatórios, fornecidas em arquivo, por meio de um conjunto de sete testes estatísticos, os quais podem ser selecionados e configurados independentemente. As conclusões de boas características das sequências e, consequentemente, dos seus respectivos geradores, é feita em razão de um nível de significância escolhido pelo usuário.

Os testes disponíveis no Sieve são apresentados brevemente a seguir:

1. **Teste da Frequência**. O foco deste teste é avaliar a proporção de 0's e 1's em uma sequência e o propósito é determinar quando este número é aproximadamente unitário, tal como esperado em uma sequência aleatória;
2. **Teste da Frequência em Blocos**. Este teste é análogo ao teste da frequência, porém baseia-se na escalabilidade dos geradores pseudo-aleatórios, checando também a frequência em subsequências da entrada, que também devem passar neste teste quando a entrada passa no teste da frequência;
3. **Teste das Corridas**. É um teste bastante utilizado cujo objetivo é a obtenção do número total de bits de um mesmo valor de forma contínua em uma sequência. Na prática, este teste detecta quando a oscilação entre 0's e 1's é muito rápida ou muito lenta;
4. **Maior Sequência de Uns**. Também utiliza a escalabilidade dos testes estatísticos de geradores pseudo-aleatórios para verificar o equilíbrio entre o comprimento das sequências de 1's consecutivos na entrada como um todo e em blocos da mesma;
5. **Teste Serial**. O teste serial analisa dois bits por vez com o propósito de determinar quando as ocorrências das subsequências 00, 01, 10 e 11 possuem aproximadamente o mesmo número;
6. **Teste Pôquer**. É uma generalização do teste da frequência e tem por objetivo determinar quando as sequências de comprimento m aparecem tal como esperado em uma sequência distribuída uniformemente;
7. **Teste da Auto-Correlação**. A autocorrelação é uma medida que informa o quanto o valor de um dos elementos da sequência de entrada é capaz de influenciar seus vizinhos. Por exemplo, o quanto a existência de um valor mais alto condiciona valores também altos de seus vizinhos na sequência. Assim, o objetivo deste teste é verificar a existência e o nível de correlações entre a sequência original e uma versão deslocada dela mesma [Guedes et al. 2009].

Tomando em conta o Sieve e os testes estatísticos que podem ser utilizados para avaliar a qualidade de geradores pseudo-aleatórios, na próxima seção serão apresentados a aplicação dos mesmos e os resultados obtidos a partir da análise de geradores pseudo-aleatórios baseados em autômatos celulares.

4. Resultados da Qualidade Estatística de Geradores Pseudo-Aleatórios Baseados em Autômatos Celulares

Para analisar a qualidade dos geradores pseudo-aleatórios baseados em autômatos celulares, considerou-se dois grupos de geradores: os geradores pseudo-aleatórios baseados em autômatos celulares elementares e aqueles baseados em autômatos celulares totalísticos.

Para a primeira classe de geradores foram consideradas 20 maneiras de inicialização diferentes, enquanto para a segunda classe 30 inicializações diferentes foram consideradas. Este número visou atender ao caráter da consistência quando se testam geradores pseudo-aleatórios, isto é, considerar a inicialização com diferentes sementes. No caso dos autômatos celulares elementares, considerou-se 20 das 256 inicializações possíveis pois almejava-se capturar principalmente sequências com padrões dos Tipos 2 e 3, que são não-triviais, visto que esse modelo não produz sequências do Tipo-4. Dentre as 30 inicializações dos autômatos celulares totalísticos, 11 têm padrão que cresce e se contrai (Tipo-4), 1 desaparece (Tipo-1) e 18 em que cresce indefinidamente (Tipo-3).

Em seguida, para cada inicialização, foram produzidos 1 milhão de bits (10^6 bits) que seriam utilizados como entrada para cada um dos testes estatísticos disponíveis na ba-

teria do Sieve. Adotou-se o nível de significância igual a 0.01, o maior valor recomendado quando se testam geradores pseudo-aleatórios.

Os resultados oriundos da bateria de testes estatísticos realizadas encontram-se descritos nas Tabelas 1 e 2, para os autômatos celulares elementares e para os totalísticos, respectivamente. As tabelas apresentam os p-valores resultantes dos respectivos testes estatísticos.

Como $\alpha = 0.01$, observa-se que para a grande maioria dos testes estatísticos realizados há rejeição da hipótese nula previamente considerada, ou seja, uma vez que p-valor $< \alpha$, é possível afirmar que as sequências em questão não são aleatórias. O único teste em que tal conclusão não pode ser afirmada é no caso da auto-correlação, no qual não foi estatisticamente possível perceber a influência de um determinado valor de bit nos seus vizinhos laterais. Este resultado possui relação direta com o tipo de padrão considerado nas saídas dos geradores.

É interessante verificar que, embora algumas sementes culminem em um nível de significância que não permite rejeitar a hipótese nula para alguns testes estatísticos, como no caso do autômato 57 elementar para os testes da frequência e da corrida e dos autômatos 1636 e 2048 totalísticos para os testes da frequência e da frequência em blocos, estes resultados constituem apenas uma porção muito pequena do total de conjunto de testes realizados, não sendo expressivos para impactar em mudança nas conclusões.

Tabela 1: Resultados dos testes estatísticos para autômatos celulares elementares.

Autômato	Tipo	Frequência	Frequência em Blocos	Corridas	Maior Sequência de Uns	Serial	Pôquer	Auto-Correlação
26	Cresce	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
30	Cresce	0.00	0.00	0.00	1.07×10^{-14}	0.00	0.00	1.00
45	Cresce	0.021	0.00	1.00	1.95×10^{-169}	0.00	0.00	1.00
57	Cresce	0.527	0.00	1.00	3.92×10^{-295}	0.00	0.00	1.00
62	Cresce	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
72	Desaparece	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
73	Cresce	0.00	0.00	0.00	1.4×10^{-137}	0.00	0.00	1.00
89	Cresce	5.6×10^{-5}	0.00	0.00	7.96×10^{-88}	0.00	0.00	1.00
91	Permanece	9.33×10^{-4}	0.00	1.00	3.97×10^{-295}	0.00	0.00	1.00
103	Permanece	5.56×10^{-231}	0.00	0.00	3.927×10^{-295}	0.00	0.00	1.00
105	Cresce	0.00	0.00	0.00	2.13×10^{-222}	0.00	0.00	1.00
109	Cresce	0.00	0.00	0.00	1.73×10^{-194}	0.00	0.00	1.00
129	Cresce	0.00	0.00	0.00	2.37×10^{-267}	0.00	0.00	1.00
136	Desaparece	0.00	0.00	0.00	1.25×10^{-238}	0.00	0.00	1.00
137	Cresce	0.00	0.00	0.00	1.84×10^{-121}	0.00	0.00	1.00
161	Cresce	0.00	0.00	0.00	3.97×10^{-274}	0.00	0.00	1.00
169	Cresce	0.00	0.00	0.00	3.92×10^{-295}	0.00	0.00	1.00
183	Cresce	0.00	0.00	0.00	$5.57 \times 10^{-}254$	0.00	0.00	1.00
214	Cresce	6.92×10^{-163}	0.00	0.00	5.39×10^{-116}	0.00	0.00	1.00
225	Cresce	0.00	0.00	0.00	3.92×10^{-295}	0.00	0.00	1.00

Como consequência destes resultados, é possível afirmar que as sequências produzidas por geradores pseudo-aleatórios construídos a partir de autômatos celulares elementares e totalísticos não são aleatórias.

Este resultando é particularmente interessante quando se verificam diversas afirmações na literatura sobre a complexidade dos padrões produzidos pela saída de autômatos celulares, classificados como sendo complexos [Wolfram 2002]. Há uma complexidade, especialmente de natureza visual destes padrões, mas tal complexidade não é suficiente para significar aleatoriedade.

Tabela 2: Resultados obtidos para autômatos celulares totalísticos.

Autômato	Tipo	Frequência	Frequência em Blocos	Corridas	Maior Sequência de Uns	Serial	Pôquer	Auto-Correlação
177	Cresce	0.00	0.00	0.00	9.25×10^{-129}	0.00	0.00	0.00
578	Cresce	2.05×10^{-119}	1.3×10^{-13}	0.00	2.89×10^{-4}	0.00	0.00	0.52
583	Cresce	0.00	0.00	0.00	0.056	0.00	0.00	1.00
600	Irregular	0.00	0.00	0.00	1.06×10^{-85}	0.00	0.00	1.00
777	Cresce	0.00	0.00	0.00	2.05×10^{-96}	0.00	0.00	0.99
843	Irregular	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
870	Irregular	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
912	Cresce	0.00	0.00	0.00	1.4×10^{-137}	0.00	0.00	1.00
993	Cresce	5.6×10^{-5}	0.00	0.00	7.96×10^{-88}	0.00	0.00	1.00
1006	Cresce	9.33×10^{-4}	0.00	1.00	3.97×10^{-295}	0.00	0.00	1.00
1020	Cresce	5.56×10^{-231}	0.00	0.00	3.927×10^{-295}	0.00	0.00	1.00
1022	Cresce	0.00	0.00	0.00	1.73×10^{-194}	0.00	0.00	1.00
1038	Cresce	0.00	0.00	0.00	2.37×10^{-267}	0.00	0.00	1.00
1041	Cresce	0.00	0.00	0.00	1.25×10^{-238}	0.00	0.00	1.00
1055	Cresce	0.00	0.00	0.00	1.84×10^{-121}	0.00	0.00	1.00
1074	Cresce	0.00	0.00	0.00	3.97×10^{-274}	0.00	0.00	1.00
1085	Cresce	0.00	0.00	0.00	3.92×10^{-295}	0.00	0.00	1.00
1092	Cresce	0.00	0.00	0.00	3.92×10^{-254}	0.00	0.00	1.00
1113	Irregular	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
1140	Irregular	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
1167	Irregular	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
1329	Irregular	0.00	0.00	0.00	3.20×10^{-92}	0.00	0.00	1.00
1572	Irregular	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
1599	Irregular	0.00	0.00	0.00	4.9×10^{-227}	0.00	0.00	1.00
1636	Irregular	0.02	1.07×10^{-71}	1.00	3.92×10^{-295}	0.00	0.00	1.00
1815	Irregular	0.00	0.00	0.00	0.00	0.00	0.00	1.00
1942	Cresce	0.00	0.00	0.00	8.26×10^{-233}	0.00	0.00	1.00
2022	Cresce	0.00	0.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
2048	Desaparece	0.01	1.00	0.00	1.26×10^{-238}	0.00	0.00	1.00
2049	Cresce	0.00	0.00	0.00	2.16×10^{-33}	0.00	0.00	1.00

Do ponto de vista prático, estes resultados impactam na escolha e utilização de geradores pseudo-aleatórios, cujos modelos baseados em autômatos celulares. Se usados em simulações, por exemplo, podem vir a comprometer a validade da mesma. Tais geradores pseudo-aleatórios devem ser evitados, devido à baixa qualidade estatística que possuem.

5. Considerações Finais

Neste trabalho foi analisada a qualidade estatística de geradores pseudo-aleatórios baseados em autômatos celulares elementares e totalísticos. Segundo a literatura, embora estruturalmente simples, este modelo produz saídas complexas.

Para realizar esta análise, foram produzidas diferentes sequências binárias a partir de diferentes inicializações destes geradores, as quais foram submetidas a uma bateria de testes estatísticos. Como resultado, foi verificado que na ampla maioria dos testes realizados não foi possível atestar aleatoriedade nas sequências produzidas por estes geradores. Apenas os testes de auto-correlação foram bem sucedidos, indicando que não é possível estabelecer uma relação entre bits vizinhos produzidos na saída do gerador.

Estes resultados mostram que a qualidade estatísticas dos geradores baseados nestas variantes de autômatos celulares é baixa e, como consequência, estes geradores devem ser evitados a fim de não comprometer as aplicações que os utilizam.

Apesar disso, os resultados oriundos neste trabalho não são definitivos sobre a exclusão de autômatos celulares como geradores pseudo-aleatórios. Como consequência, abre-se a possibilidade de novas maneiras de explorar este modelo de computação para

tais geradores: mudando a maneira como as sequências são produzidas, utilizando regras mais complexas, combinando com outros geradores e algoritmos, etc.

Em trabalhos futuros, almeja-se prosseguir no estudo deste modelo de computação. Primeiramente, almeja-se considerar a qualidade estatística de geradores pseudo-aleatórios baseados em modelos de autômatos celulares mais complexos, tais como os bidimensionais. Além disso, almeja-se investigar outras maneiras de produzir sequências pseudo-aleatórias a partir deste geradores, especialmente visando descartar padrões dos Tipos 1 e 2.

Agradecimentos

As autoras agradecem o apoio financeiro provido pela Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM). A autora Nicoli Pinheiro de Araújo é bolsista de iniciação científica do Programa de Apoio à Iniciação Científica da UEA/FAPEAM Edição 2014-2015.

Referências

Bardell, P. H. (1990). Analysis of cellular automata used as pseudorandom pattern generators. In *International Test Conference*, pages 762–768.

Chen, Q., Mynett, A., and Minns, A. (2002). Application of cellular automata to modelling competitive growths of two underwater species chara aspera and potamogeton pectinatus in lake veluwe. *Ecological Modelling*, 147(3):253–265.

Gentle, J. E. (2003). *Random Number Generation and Monte Carlo Methods*. Springer, Estados Unidos.

Guan, S.-U. and Tan, S. (2004). Pseudorandom number generation with self-programmable cellular automata. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 23(7):1095–1101.

Guedes, E. B., dos Santos, G. O., and de Assis, F. M. (2009). Sieve – statistical analysis of random and pseudo random number generators. Disponível em `http://sieve.googlecode.com`. Último acesso em Setembro de 2015.

Kang, B.-H., Lee, D.-H., , and Hong, C.-P. (2008). *Novel Algorithms and Techniques in Telecommunications, Automation and Industrial Electronics*, chapter Pseudorandom Number Generation Using Cellular Automata, pages 401–404. Springer.

Knuth, D. E. (1998). *The Art of Computer Programming – Volume 2 – Seminumerical Algorithms*. Addison-Wesley Publishing Company.

L'Ecuyer, P. (1992). Testing random number generators. In *Proceedings of the 1992 Winter Simulation Conference*, pages 305–313, Virginia, Estados Unidos. Association for Computing Machinery.

Marsaglia, G. The Marsaglia Random Number CDROM, including the DIEHARD.

Menezes, A., van Oorschot, P., and Vanstone, S. (1996). *Handbook of Applied Cryptography*. CRC Press.

Navid, A. H. F. and Aghababa, A. B. *Emerging Applications of Cellular Automata*. In-Tech, Estados Unidos.

Popovici, A. and Popovici, D. (2002). Cellular automata in image processing. In *Fifteenth International Symposium on Mathematical Theory of Networks and Systems*, volume 1.

Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Leveson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., and Vo, S. (2008). A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Technical report, National Institute of Standards and Technology – Technology Administration – U. S. Department of Commerce, Estados Unidos.

Salcido, A. (2013). *Emerging Applications of Cellular Automata*. InTech, Estados Unidos.

Schiff, J. L. (2010). *Cellular Automata: A Discrete View of the World*. Wiley-Interscience, 1 edition.

Williams, C. P. and Clearwater, S. H. (1997). *Explorations in quantum computing*. Springer-Verlag TELOS.

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature*, 311(5985):419–424.

Wolfram, S. (1986). Random sequence generation by cellular automata. *Advances In Applied Mathematics*, 7:123–169.

Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media, Champaign, IL.

Ye, F., Chen, Q., and Li, R. (2010). Modelling the riparian vegetation evolution due to flow regulation of lijiang river by unstructured cellular automata. *Ecological informatics*, 5(2):108–114.