Sequência Pea Pattern

André Kowacs
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo
andrekowacs@gmail.com

Resumo

O artigo apresenta uma família de sequências baseada na "Look and Say" estudada pelo matemático John Conway. São estudados padrões que se repetem na sequência, pontos fixos, bem como será apresentada uma tabela listando todos os pontos fixos para a sequência sobre determinados alfabetos, ou seja, conjunto de simbolos.

1. Introdução

Considere a seguinte sequência numérica:

$$1, 11, 21, 1211, 1231, 131221, 132231, 232221, 134211, 14131231, ...$$

Você consegue adivinhar o próximo número? A sequência camuflada (ou "Pea Pattern" em inglês) é definida da seguinte forma: o termo seguinte é contagem, em ordem decrescente dos algarismos do termo anterior. Por exemplo, depois de 1231 vem 131221 pois há 1 "3", 1 "2" e 2 "1"s em 1231. Ainda não entendeu? Tente ler em voz alta: 1231 tem um três, um dois e dois um. Vamos continuar a sequência anterior: 14231241,24132231,14233221,14233221,....
Neste exemplo vemos que a sequência "converge" para o termo 14233221. Será que isso sempre acontece? Existem outros termos nos quais a sequência "estaciona"? Note que escolhemos "contar"os números em base 10, porém se restringirmos os números que trabalhamos podemos fazer essa contagem em outras bases numéricas também. Será que em alguma base essas sequências tem alguma propriedade especial? Estas questões motivaram a escrita deste artigo, o qual procura respondê-las.

Vale notar que esta família de sequências é parecida com a estudada pelo matemático John Conway, a família das Look and Say, cuja diferença é que se conta o número de algarismos iguais juntos apenas. Para mais informações sobre esta, veja a referência [1].

2. Considerações iniciais e notações

Inicialmente vamos definir os conceitos básicos usados no artigo. Note que não podemos tratar os termos das sequências como números inteiros simplesmente, pois precisaremos contar seus algarismos. Note também que dependendo da base na qual escreveremos os números os algarismos usados irão mudar, por exemplo, se resolvermos escrever a sequência na base 2 (binária) e começarmos com 0, os próximos termos serão: $10, 1110, 11110, 100110, ...$ De modo que precisaremos introduzir a noção de "alfabeto".
Dizemos que um conjunto Σ composto por símbolos distintos é um *alfabeto*. Uma *palavra* sobre

o alfabeto Σ é uma lista de símbolos de Σ, justapostos pela operação de concatenação "$||$", que será, em geral, omitida. Por exemplo, dado o alfabeto $\Sigma = \{0, 1, 2\}$ um exemplo de palavra é $w = 1223$. Denotamos por Σ^* o conjunto de todas as palavras finitas sobre o alfabeto Σ. Por exemplo, dado $\Sigma = \{0, 1\}$, temos que $10, 11, 10101, 1111, ... \in \Sigma^*$. Notamos que para quando a letra $0 \in \Sigma$, consideraremos pertencentes a Σ^* apenas as palavras cuja primeira letra é diferente de 0. Por exemplo, se $\Sigma = \{0, 1, 3\}$ então 011301 não pertence a Σ^*, mas 11301 sim.
Por fim, dado um número x real, denotaremos por $\lceil x \rceil$ o menor inteiro maior ou igual a x e $\lfloor x \rfloor$ o maior inteiro menor ou igual a x.

Observa-se que ao longo do artigo se abusará da bijeção natural existente entre a representação em base numérica k de $\mathbb{N}$ e o conjunto de palavras (que não começam com 0) Σ_k^* sobre o alfabeto $\Sigma_k = \{0, 1, ..., k-1\}$. Isto é, a bijeção que leva, por exemplo, a palavra $w = 12 \in \Sigma_{10}^*$ no número representado por $n = 12$ na base numérica $k = 10$. Assim podemos estender, por exemplo, a relação de ordem de $\mathbb{Z}$ para Σ_k^*, ao estabelecer que $w_1 \leq w_2$ se o inteiro associado a w_1 for menor ou igual ao inteiro associado a w_2.

Vamos denotar por $|x|_i$ o número de letras $i \in \Sigma$ na palavra $x \in \Sigma^*$. Ainda, $|x|$ denotará o comprimento da palavra $x \in \Sigma_k^*$. Por exemplo, seja $\Sigma = \{A, R, S\}$ e $w = ARARA \in \Sigma^*$. Então $|w|_A = 3, |w|_R = 2, |w|_S = 0$ e $|w| = 5$.

Agora considere um alfabeto Σ e uma palavra $x \in \Sigma^*$. Definimos

$$Im(x) = \{l \in \Sigma | |x|_l > 0\}$$

Por exemplo, seja $1123 \in \Sigma_1^* 0$ então $Im(1143) = \{1, 3, 4\}$.
Finalmente estamos prontos para definir formalmente a família de sequências Pea Pattern. A fim de esclarecer como será definida, vejamos um exemplo. Dado $x_0 = 123 \in \Sigma_{10}^*$, temos:

$$|x_0|_0 = 0, |x_0|_1 = 1, |x_0|_2 = 1, |x_0|_3 = 1, |x_0|_4 = 0, ..., |x_0|_9 = 0$$

Logo $Im(x) = \{1, 2, 3\}$. De modo que, ordenando os elementos de $Im(x)$ temos: $b_1 = 3, b_2 = 2, b_3 = 1$ obtemos o próximo termo da sequência: $x_1 = |x_0|_{b_1} || b_1 ||| x_0|_{b_2} || b_2 ||| x_0|_{b_3} || b_3$ ou simplesmente: $x_1 = 131211$.
A família de sequências é então formalmente definida por:
Dado $x_0 \in \Sigma_k^*$, para cada $n \in \mathbb{N}$, tome $b_1, ..., b_r \in \Sigma$ tais que $b_i \in Im(x_n), \forall i$, $b_1 > b_2 > ... > b_r$ e $\bigcup_1^r b_i = Im(x_n)$. Definimos a sequência $(x_n)_n \in \Sigma_k^*$ recursivamente pela expressão:

$$x_{n+1} = (|x_n|_{b_1})_k \, b_1 \, (|x_n|_{b_2})_k \, b_2 \ldots (|x_n|_{b_r})_k \, b_r$$

Tal que $(|x_n|_{b_j})_k$, é o número$|x_n|_{b_j}$ na base k.

De volta ao exemplo anterior, computando o restante de termos da sequência, notamos que esta estabiliza na a palavra $x = 14233221 \in \Sigma_{10}^*$, isto é, se $x_N = 14233221$, então $x_{N+k} = 14233221$, para todo $k \in \mathbb{N}$. Chamaremos palavras com essa propriedade de *ponto fixo* do nosso universo de palavras. Além disso, se $x_N = x_{N+p}$, para algum $p \in \mathbb{N}$, tomamos o menor p para que isto vale, dizemos que x_N é um *p-ciclo* em Σ_k^*.

3. Pontos fixos e ciclos

Note que quando começamos com $x_0 = 1$, obtemos o ponto fixo 14233221. Essa ocorrência da sequência de estabilizar em um ponto fixo também acontece com outros valores de x_0. A pergunta

natural que surge então é se para dado qualquer $x_0 \in \Sigma_k^*$ o sistema sempre se estabiliza? A resposta é sim. Veremos através do seguinte teorema, que usa do fato de a representação numérica de um número ter, em geral, comprimento menor que o número em si.
Mas antes, apresento dois lemas que serão usados na sua demonstração.

Lema 1. *Sejam $x \in \Sigma_k^*$. Então $|x| = \lfloor log_k x \rfloor + 1$.*

Demonstração. Seja $x \in \Sigma_k^*$, tal que $|x| = n+1$. Então $x = a_n a_{n-1} ... a_1 a_0$, com $0 \leq a_i \leq k-1, \forall i$ e $1 \leq a_n \leq k-1$. Ao mesmo tempo, $x = a_n * k^n + ... + a_0 * k^0$. Assim,

$$\lfloor \log_k x \rfloor = \lfloor \log_k (a_n * k^n + ... + a_0 * k^0) \rfloor < \lfloor \log_k k^{n+1} \rfloor = n+1$$

E ainda:

$$\lfloor \log_k x \rfloor = \lfloor \log_k (a_n * k^n + ... + a_0 * k^0) \rfloor \geq \lfloor \log_k k^n \rfloor = n$$

Portanto $\lfloor log_k x \rfloor = n$ e $\lfloor log_k x \rfloor + 1 = n+1 = |x|$ □

Lema 2. *Sejam $k, r \in \mathbb{N}$, $k \geq 2$. Então vale a estimativa:*

$$\lfloor \log_k r \rfloor \leq r/k^2 + 1$$

Demonstração. Omitimos a demonstração dessa identidade uma vez que consiste num simples porém extenso trabalho numérico. Ao leitor interessado, basta reparar que $\lfloor \log_k r \rfloor \leq \log_k r$ e avaliar o mínimo da função $f(x,k) = x/k^2 + 1 - \log_k x$, para $k \geq 2$.

□

Enfim temos as ferramentas necessárias para provar o teorema principal. Independente do alfabeto Σ_k e da palavra inicial x_0, a sequência Pea Pattern estabiliza com um ponto fixo ou um p-ciclo.

Teorema 3. *Para todo $x_0 \in \Sigma_k$, a sequência definida por $x_{n+1} = \mathcal{P}_k(x_n)$ converge para um ponto fixo ou ciclo.*

Demonstração. Inicialmente, note que o número de palavras com no máximo uma quantidade de letras $l \in \mathbb{N}$ é finito. De fato, como temos k letras em Σ_k, existem $\frac{k^{n+1}-k}{k-1}$ palavras com até n letras em Σ_k^* (Soma de uma PG). Logo se mostrarmos que o número de letras dos termos da sequência é limitado para n suficientemente grande, concluímos que haverá um limite para a quantidade de palavras diferentes.
Seja $x_n \in \Sigma_k^*$. Então, pela definição, temos:

$$\begin{aligned} x_{n+1} &= (|x_n|_{b_1})_k\, b_1 ... (|x_n|_{b_r})_k b_r \\ |x_{n+1}| &= |(|x_n|_{b_1})_k| + 1 + ... + |(|x_n|_{b_r})_k| + 1 \\ &\leq |(|x_n|_{k-1})_k| + 1 + ... + |(|x_n|_0)_k| + 1 \\ &\leq |(|x_n|_{k-1})_k| + ... + |(|x_n|_0)_k| + k \end{aligned}$$

Mas note que $|(|x_n|_{k-1})_k| + ... + |(|x_n|_0)_k| \leq k|(|x_n|)_k|$, logo:

$$|x_{n+1}| \leq k|(|x_n|)_k| + k$$

Aplicando o Lema 1, temos então que:

$$|x_{n+1}| \leq k(\lfloor \log_k |x_n| \rfloor + 2) \tag{1}$$

Aplicando a desigualdade do Lema 2, temos então, de (1) que:

$$|x_{n+1}| \leq \frac{|x_n|}{k} + 3k$$

Logo, para $|x_n| > \frac{3k^2}{k-1}$, temos que:

$$|x_{n+1}| < |x_n|$$

De fato, pois note que se

$$|x_n| > \frac{3k^2)}{k-1}$$

Então

$$k-1 > \frac{3k^2}{|x_n|}$$

De onde segue que

$$|x_{n+1}| \leq \frac{|x_n|}{k} + 3k = |x_n|\left(\frac{1}{k} + \frac{3k}{|x_n|}\right) =$$
$$= |x_n|\left(\frac{1}{k^2}\left(1 + \frac{3k^2}{|x_n|}\right)\right) < |x_n|\left(\frac{1}{k}(1+k-1)\right) = |x_n|$$

Além disso, para $|x_n| \leq \frac{3k^2}{k-1}$, implica que

$$|x_{n+1}| \leq \frac{|x_n|}{k} + 3k \leq \frac{3k}{k-1} + 3k = \frac{3k^2}{k-1} := \alpha_k$$

Logo, seja $m = \max\{|x_0|, \lceil \alpha_k \rceil\}$. Então $|x_n| \leq m \quad \forall n \in \mathbb{N}$, isto pois se $|x_0| > \frac{3k^2}{k-1}$, temos que $|x_n| < |x_0|$, $\forall n$, e se $|x_0| \leq \lceil \alpha_k \rceil$, temos que $|x_n| < \lceil \alpha_k \rceil$, $\forall n$. □

Mais ainda, o resultado diz que $|x_n| \leq \lceil \alpha_k \rceil$ para todo $n > N \in \mathbb{N}$ suficientemente grande.

Assim temos que todos os pontos fixos e ciclos num alfabeto Σ_k^* tem comprimento menor ou igual a $\lceil \alpha_k \rceil$. Note que como existem k letras em Σ_k, o número de palavras com comprimento l é dado por k^l. Assim o número de palavra com comprimento menor ou igual a $\lceil \alpha_k \rceil$ é dado por:

$$n = k^1 + k^2 + ... + k^{\lceil \alpha_k \rceil} = \frac{k^{\lceil \alpha_k \rceil + 1} - k}{k-1}$$

Fica simples então mostrar que para $k = 2$ os únicos pontos fixos são $x' = 111$ e $x'' = 1001110$, por exemplo, pois basta testar $n = 8190$ palavras. Porém note que n cresce exponencialmente com k. O grande tempo de processamento para verificar todas essas palavras motivou o segundo teorema, que dá mais condições necessárias para uma palavra ser ponto fixo.

Teorema 4. *Seja $\overline{x} \in \Sigma_k^*$ ponto fixo de uma sequência da forma descrita acima, tal que:*

$$\overline{x} = a_{(n_1,1)}a_{(n_1-1,1)}...a_{(0,1)}b_1a_{(n_2,2)}...a_{(0,2)}b_2...a_{(n_r,r)}...a_{(0,r)}b_r$$

Com $a_{(i,j)} \in \Sigma_k$*,* $\forall i < n_j$ *e* $a_{(n_j,j)} \in \Sigma_k - \{0\}$ *e* $0 \leq n_j$*;* $1 \leq j \leq r$*.*
Então:

$$\sum_{j=1}^{r} n_j \geq \sum_{j=1}^{r} k^{n_j} - 2r$$

Demonstração. Inicialmente, note que:

$$|\overline{x}| = (n_1 + 1) + 1 + (n_2 + 1) + 1 + ... + (n_r + 1) + 1 = \sum_{j=1}^{r} n_j + 2r \tag{2}$$

Além disso, como $\overline{x}$ é ponto fixo, logo:

$$\begin{gathered} a_{(n_1,1)} \cdot k^{n_1-1} + ... + a_{(0,1)} \cdot k^0 = |\overline{x}|_{b_1} \\ a_{(n_2,2)} \cdot k^{n_2-1} + ... + a_{(0,2)} \cdot k^0 = |\overline{x}|_{b_2} \\ \vdots \\ a_{(n_r,r)} \cdot k^{n_r-1} + ... + a_{(0,r)} \cdot k^0 = |\overline{x}|_{b_r} \\ \therefore \\ \sum_{j=1}^{r} \left(a_{(n_j,j)} \cdot k^{n_j} + ... + a_{(0,j)} \cdot k^0 \right) = |\overline{x}| \end{gathered}$$

Como $0 \leq a_{(i,j)}$, temos:

$$\sum_{j=1}^{r} \left(a_{(n_j,j)} \cdot k^{n_j} \right) \leq |\overline{x}|$$

Como $1 \leq a_{(n_j,j)}$, temos:

$$\sum_{j=1}^{r} k^{n_j} \leq |\overline{x}|$$

Mas então, a equação (2) implica:

$$\sum_{j=1}^{r} k^{n_j} \leq \sum_{j=1}^{r} n_j + 2r$$

Logo,

$$\sum_{j=1}^{r} n_j \geq \sum_{j=1}^{r} k^{n_j} \ - 2r$$

□

Na verdade, esse teorema é um caso específico do caso mais geral tratado em seguida.
Se denotarmos por $\overline{x}_i$ para $i = 1, ..., p$ as palavras que formam um p-ciclo, tais que, para $i = 1, ..., p$:

$$\overline{x}_i = a_{(n_{(1,i)},1,i)} a_{(n_{(1,i)}-1,1,i)} ... a_{(0,1,i)} b_1 a_{(n_{(2,i)},2,i)} ... a_{(0,2,i)} b_2 ... a_{(n_{(r,i)},r_i,i)} ... a_{(0,r,i)} b_r$$

Teorema 5. *Sejam* $\overline{x}_i$ *as palavras que formam um* p*-ciclo em* Σ_k^**. Então:*

$$\sum_{i=1}^{p} \sum_{j=1}^{r} n_{(j,i)} \geq \sum_{i=1}^{p} \sum_{j=1}^{r} k^{n_{(j,i)}} \ - 2pr$$

Demonstração. Segue similar à do Teorema 2. □

Dos três teoremas é fácil verificar computacionalmente o seguinte corolário:

Corolário 1. *A sequência da forma* $x_{n+1} = \mathcal{P}_2(x_n)$*, converge para o ponto fixo:* $\overline{x} = 1001110$*, para todo* $x_0 \in \Sigma_2^*$*, exceto para* $x_0 = \overline{x}' = 111$.

Ou seja, não existem ciclos com $p \geq 2$ na base 2. O mesmo já não é verdade da base 3 em diante.

4. Conclusão

Apesar destes teoremas diminuírem o número de palavras a serem testadas para pontos fixos ou ciclos, a implementação computacional é mais complicada, o que dificulta a verificação de ciclos e pontos fixos para bases maiores, ficando em aberto, ainda, o problema de achar uma descrição eficiente para os pontos fixos e ciclos da sequência numa dada base. Também parece difícil determinar, dado x_0, qual ponto fixo ou ciclo a sequência irá convergir, tendo em vista seu comportamento caótico. Uma análise mais profunda neste ponto seria crucial para o entendimento e maior elaboração desta teoria.

Pontos fixos para Σ_k^*, com $k =$				
2	3	4	5	6
111	22	22	22	22
1001110	11110	1211110	14233221	14233221
	12111	1311110	14331231	14331231
	101100	1312111	14333110	14333110
	1022120	23322110	23322110	15143331
	2211110	33123110	33123110	15233221
	22101100	132211110	131211110	15331231
			141211110	15333110
			141311110	23322110
			141312111	33123110
			1433223110	1433223110
			14132211110	1514332231
				1533223110
				14131211110
				15131211110
				15141311110
				15141312111
				1514132211110

Alguns pontos fixos para Σ_k^*, com $k =$			
7	8	9	10
22	22	22	22
14233221	14233221	14233221	14233221
14331231	14331231	14331231	14331231
14333110	14333110	14333110	14333110
15143331	15143331	15143331	15143331
15233221	15233221	15233221	15233221
15331231	15331231	15331231	15331231
15333110	15333110	15333110	15333110
16143331	16143331	16143331	16143331
16153331	16153331	16153331	16153331
16233221	16233221	16233221	16233221
16331231	16331231	16331231	16331231
16333110	16333110	16333110	16333110
23322110	17143331	17143331	17143331
33123110	17153331	17153331	17153331
1433223110	17163331	17163331	17163331
1514332231	17233221	17233221	17233221
1533223110	17331231	17331231	17331231
1614332231	17333110	17333110	17333110
1615332231	23322110	18143331	18143331
1633223110	33123110	18153331	18153331
	1433223110	18163331	18163331
	1514332231	18173331	18173331
	1533223110	18233221	18233221
	1614332231	18331231	18331231
	1615332231	18333110	18333110
	1633223110	23322110	19143331
	1714332231	33123110	19153331
	1715332231	1433223110	19163331
	1716332231	1514332231	19173331
	1733223110	1533223110	19183331
	16152423324110	1614332231	19233221
	17152423324110	1615332231	19331231
	17161524233241	1633223110	19333110
	17162423324110	1714332231	33123110
	161514131211110	1715332231	1433223110
	171514131211110	1716332231	1514332231
	171614131211110	1733223110	1533223110
	171615131211110	1814332231	1614332231
	171615141211110	1815332231	1615332231
	171615141311110	1816332231	1633223110
	171615141312111	1817332231	1714332231
	1716251423325110	1833223110	1715332231
	17161514132211110	16152423324110	1716332231

Alguns pontos fixos para Σ_k^*, com $k =$	
9	10
17152423324110	1733223110
17161524233241	1814332231
17162423324110	1815332231
18152423324110	1816332231
18161524233241	1817332231
18162423324110	1833223110
18171524233241	1914332231
18171624233241	1915332231
18172423324110	1916332231
1716251423325110	1917332231
1816251423325110	1918332231
1817162514233251	1933223110
1817162523325110	16152423324110
1817251423325110	17152423324110
171615141312111110	17161524233241
181615141312111110	17162423324110
181715141312111110	18152423324110
181716141312111110	18161524233241
181716151312111110	18162423324110
181716151412111110	18171524233241
181716151413111110	18171624233241
181716151413121111	18172423324110
181716251433325110	19152423324110
181726151423326110	19161524233241
18171615141322111110	19162423324110
	19171524233241
	19171624233241
	19172423324110
	19181524233241
	19181624233241
	19181724233241
	19182423324110
	1716251423325110
	1816251423325110
	1817162514233251
	1817162523325110
	1817251423325110
	1916251423325110
	1917162514233251

Alguns pontos fixos para Σ_k^*, com $k =$ 10
1917162523325110
1917251423325110
1918162514233251
1918162523325110
1918171625233251
1918172514233251
1918172523325110
1918251423325110
181726151423326110
191726151423326110
191817261423326110
191817261514233261
191817261523326110
191826151423326110
181716151413121111 0
1917161514131211110
1918161514131211110
1918171514131211110
1918171614131211110
1918171615131211110
1918171615141211110
1918171615141311110
1918171615141312111
19182716151423327110
191817161514132211110

Referências

[1] Conway, J. H. "The Weird and Wonderful Chemistry of Audioactive Decay."Eureka 46, 5-18, 1986.

[2] Jean-Paul Allouche and Jeffrey Shallit. 2003. Automatic Sequences: Theory, Applications, Generalizations. Cambridge University Press, New York, NY, USA.

[3] OEIS Foundation Inc. (2017), The On-Line Encyclopedia of Integer Sequences, https://oeis.org/A005150