Regular Sequences

Eric Rowland

School of Computer Science
University of Waterloo, Ontario, Canada

September 5, 2011
Outline

1 Motivation and basic properties

2 Sampler platter

3 Relationships to other classes of sequences
Everyone’s favorite sequence

Thue–Morse sequence:

\[ a(n) = \begin{cases} 
0 & \text{if the binary representation of } n \text{ has an even number of 1s} \\
1 & \text{if the binary representation of } n \text{ has an odd number of 1s.} 
\end{cases} \]

For \( n \geq 0 \), the Thue–Morse sequence is

\[ 0 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ \cdots \ . \]

Rediscovered several times as an infinite cube-free word on \( \{0, 1\} \).

\[
\begin{align*}
a(2n + 0) &= a(n) \\
a(2n + 1) &= 1 - a(n)
\end{align*}
\]
Let $\nu_k(n)$ be the exponent of the largest power of $k$ dividing $n$.

For $n \geq 0$, the “ruler sequence” $\nu_2(n + 1)$ is

$$0 \ 1 \ 0 \ 2 \ 0 \ 1 \ 0 \ 3 \ 0 \ 1 \ 0 \ 2 \ 0 \ 1 \ 0 \ 4 \ 0 \ 1 \ 0 \ 2 \ 0 \ 1 \ 0 \ 3 \ 0 \ 1 \ 0 \ 2 \ 0 \ 1 \ 0 \ 5 \ \cdots$$

$$\nu_2(2n + 0) = 1 + \nu_2(n)$$
$$\nu_2(2n + 1) = 0$$
Let $a(n) = |\{0 \leq m \leq n : \binom{n}{m} \not\equiv 0 \mod 8\}|$.

$a(2n + 1) = 2a(n)$

$a(4n + 0) = a(2n)$

$a(8n + 2) = -2a(n) + 2a(2n) + a(4n + 2)$

$a(8n + 6) = 2a(4n + 2)$
Convention:
We index sequences starting at $n = 0$.

**Definition (Allouche & Shallit 1992)**

Let $k \geq 2$ be an integer. An integer sequence $a(n)$ is $k$-regular if the $\mathbb{Z}$-module generated by the set of subsequences

$$\{ a(k^e n + i) : e \geq 0, 0 \leq i \leq k^e - 1 \}$$

is finitely generated.

We can take the generators to be elements of this set. Every $a(k^e n + i)$ is a linear combination of the generators.

In particular, $a(k^e (kn + j) + i)$ is a linear combination of the generators, which gives a finite set of recurrences that determine $a(n)$.
Homogenization

For the Thue–Morse sequence:

\[ a(2n + 0) = a(n) \]
\[ a(2n + 1) = 1 - a(n) \]

But we can homogenize:

\[ a(2n) = a(n) \]
\[ a(2n + 1) = a(2n + 1) \]
\[ a(4n + 1) = a(2n + 1) \]
\[ a(4n + 3) = a(n) \]

So \( a(n) \) and \( a(2n + 1) \) generate the \( \mathbb{Z} \)-module, and we have written \( a(2n + 0), a(2n + 1), a(2(2n + 0) + 1), a(2(2n + 1) + 1) \) as linear combinations of the generators.
Basic properties

Regular sequences inherit self-similarity from base-

The \( n \)th term \( a(n) \) can be computed quickly — using \( O(\log n) \) additions and multiplications.

The set of \( k \)-regular sequences is closed under . . .

- termwise addition
- termwise multiplication
- multiplication as power series
- shifting \( (b(n) = a(n + 1)) \)
- modifying finitely many terms
Outline

1. Motivation and basic properties

2. Sampler platter

3. Relationships to other classes of sequences
More examples

Regular sequences are everywhere...

- The length \( a(n) \) of the base-\( k \) representation of \( n + 1 \) is a \( k \)-regular sequence:

  \[
  a(kn + i) = 1 + a(n).
  \]

- The number of comparisons \( a(n) \) required to sort a list of length \( n \) using merge sort is

  \[
  0 \ 0 \ 1 \ 3 \ 5 \ 8 \ 11 \ 14 \ 17 \ 21 \ 25 \ 29 \ 33 \ 37 \ 41 \ 45 \ \ldots.
  \]

  This sequence satisfies

  \[
  a(n) = a(\left\lfloor \frac{n}{2} \right\rfloor) + a(\left\lceil \frac{n}{2} \right\rceil) + n - 1
  \]

  and is 2-regular.
The coordinates \((x(n), y(n))\) of paperfolding curves are 2-regular.
\( \nu_k(n + 1) \) is \( k \)-regular:

\[
a(kn + k - 1) = 1 + a(n)
\]

\[
a(kn + i) = 0 \quad \text{if } i \neq k - 1.
\]

Bell 2007:
If \( f(x) \) is a polynomial, \( \nu_p(f(n)) \) is \( p \)-regular if and only if \( f(x) \) factors as

(product of linear polynomials over \( \mathbb{Q} \)) \( \cdot \) (polynomial with no roots in \( \mathbb{Z}_p \)).

\( \nu_p(n!) \) is \( p \)-regular.

Closure properties imply that

\[
\nu_p(C_n) = \nu_p((2n)! - 2\nu_p(n!) - \nu_p(n+1)
\]

is \( p \)-regular.
$p$-adic valuations of integer sequences

Medina–Rowland 2009:
\[ \nu_p(F_n) \text{ is } p\text{-regular.} \]

The Motzkin numbers $M_n$ satisfy

\[ (n + 2)M_n - (2n + 1)M_{n-1} - 3(n - 1)M_{n-2} = 0. \]

**Conjecture**

If $p = 2$ or $p = 5$, then $\nu_p(M_n)$ is $p$-regular.

**Open question**

Given a polynomial-recursive sequence $f(n)$, for which primes is $\nu_p(f(n))$ $p$-regular?
The sequence of integers expressible as a sum of distinct powers of 3 is 2-regular:

\[ a(2n) = 3a(n) \]
\[ a(4n + 1) = 6a(n) + a(2n + 1) \]
\[ a(4n + 3) = -3a(n) + 4a(2n + 1) \]

The sequence of integers whose binary representations contain an even number of 1s is 2-regular:

\[ 0 \ 1 \ 3 \ 4 \ 9 \ 10 \ 12 \ 13 \ 27 \ 28 \ 30 \ 31 \ 36 \ 37 \ 39 \ 40 \ \cdots \]

Let \( |n|_w \) be the number of occurrences of \( w \) in the base-\( k \) representation of \( n \). For every word \( w \), \( |n|_w \) is \( k \)-regular.
Let $a_p^\alpha(n) = |\{0 \leq m \leq n : \binom{n}{m} \not\equiv 0 \mod p^\alpha\}|$.

• Glaisher 1899:
  $$a_2(n) = 2^{\mid n \mid_1}.$$

• Fine 1947:
  $$a_p(n) = \prod_{i=0}^{l} (n_i + 1),$$
  where $n = n_l \cdots n_1 n_0$ in base $p$.
  
  For example, $a_5(n) = 2^{\mid n \mid_1} 3^{\mid n \mid_2} 4^{\mid n \mid_3} 5^{\mid n \mid_4}$.

It follows that $a_p(n)$ is $p$-regular.
Nonzero binomial coefficients

Rowland 2011:
Algorithm for obtaining a symbolic expression in $n$ for $a_{p^\alpha}$
Nonzero binomial coefficients

Higher powers of 2:

$$a_8(n) = 2^{|n|_1} \left( 1 + \frac{1}{8} |n|_{10}^2 + \frac{3}{8} |n|_{10} + |n|_{100} + \frac{1}{4} |n|_{110} \right)$$

$$\frac{a_{16}(n)}{2^{|n|_1}} = 1 + \frac{5}{12} |n|_{10} + \frac{1}{2} |n|_{100} + \frac{1}{8} |n|_{110}$$

$$+ 2 |n|_{1000} + \frac{1}{2} |n|_{1010} + \frac{1}{2} |n|_{1100} + \frac{1}{8} |n|_{1110} + \frac{1}{16} |n|_{10}^2$$

$$+ \frac{1}{2} |n|_{10} |n|_{100} + \frac{1}{8} |n|_{10} |n|_{110} + \frac{1}{48} |n|_{10}^3$$
Powers of polynomials

If \( f(x) \in \mathbb{F}_{p^\alpha}[x] \), how many nonzero terms are there in \( f(x)^n \)? Such a sequence has an interpretation as counting cells in a cellular automaton. Here is \((x^d + x + 1)^n\) over \(\mathbb{F}_2\) for \(d = 2, 3, 4\):

Amdeberhan–Stanley \(\sim 2008\):
Let \( f(x_1, \ldots, x_m) \in \mathbb{F}_{p^\alpha}[x_1, \ldots, x_m] \). The number \( a(n) \) of nonzero terms in the expanded form of \( f(x_1, \ldots, x_m)^n \) is \(p\)-regular.
Another kind of self-similarity

Here is a cellular automaton that grows like $\sqrt{n}$:

The length of row $n$ is 2-regular.

The number of black cells on row $n$ is 2-regular.
What is the lexicographically least square-free word on $\mathbb{Z}_{\geq 0}$?

$$0 1 0 2 0 1 0 3 0 1 0 2 0 1 0 4 0 1 0 2 0 1 0 3 0 1 0 2 0 1 0 5 \ldots$$

The $n$th term is $\nu_2(n + 1)$.

The lexicographically least $k$-power-free word is given by $\nu_k(n + 1)$.

$k = 3$:

$$0 0 1 0 0 1 0 0 2 0 0 1 0 0 1 0 0 2 0 0 1 0 0 1 0 0 3 0 0 1 0 0 \ldots$$

$k = 4$:

$$0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 2 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 2 \ldots$$
If \( w = w_1 w_2 \cdots w_l \) is a length-\( l \) word and \( r \in \mathbb{Q}_{\geq 0} \) such that \( r \cdot l \in \mathbb{Z} \), let

\[
w^r = w^{\lfloor r \rfloor} w_1 w_2 \cdots w_{l \cdot (r - \lfloor r \rfloor)}
\]

be the word consisting of repeated copies of \( w \) truncated at \( rl \) letters.

For example...

\[(\text{deci})^{3/2} = \text{decide}\]

\[(\text{raisonne})^{3/2} = \text{raisonnerais}\]

\[(\text{schuli})^{3/2} = \text{schulisch}\]
Lexicographically extremal words avoiding a pattern

What is the lexicographically least word on $\mathbb{Z}_{\geq 0}$ avoiding $3/2$-powers?

0 0 1 1 0 2 1 0 0 1 1 2 0 0 1 1 0 3 1 0 0 1 1 3 0 0 1 1 0 2 1 0 0 1 1 4 · · ·

Rowland–Shallit 2011:
This sequence is 6-regular.

Open question

*When are such sequences $k$-regular, and for what value of $k$?*
Outline

1. Motivation and basic properties
2. Sampler platter
3. Relationships to other classes of sequences
The companion matrix of a constant-recursive sequence \( a(n) \) satisfies

\[
M \cdot \begin{bmatrix}
a(n) \\
a(n+1) \\
\vdots \\
a(n+r-1)
\end{bmatrix} = \begin{bmatrix}
a(n+1) \\
a(n+2) \\
\vdots \\
a(n+r)
\end{bmatrix}.
\]

For example,

\[
\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} F_n \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} F_{n+1} \\ F_n + F_{n+1} \end{bmatrix} = \begin{bmatrix} F_{n+1} \\ F_{n+2} \end{bmatrix}.
\]

So

\[
F_n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\]

In general \( a(n) = \lambda M^n \kappa \).
Matrix formulation

Take \( r \) generators \( a_1(n), \ldots, a_r(n) \) of a \( k \)-regular sequence. Each \( a_j(kn + i) \) is a linear combination of the \( r \) generators.

Encode the coefficients in \( r \times r \) matrices \( M_0, M_1, \ldots, M_{k-1} \). Then if \( n = n_l \cdots n_1 n_0 \) in base \( k \), then

\[
a(n) = \lambda M_{n_l} \cdots M_{n_1} M_{n_0} \kappa.
\]

Again consider the Thue–Morse sequence; generators \( a(n), a(2n + 1) \).

\[
a(2n) = 1 \cdot a(n) + 0 \cdot a(2n + 1)
\]
\[
a(2n + 1) = 0 \cdot a(n) + 1 \cdot a(2n + 1)
\]
\[
a(2(2n + 0) + 1) = 0 \cdot a(n) + 1 \cdot a(2n + 1)
\]
\[
a(2(2n + 1) + 1) = 1 \cdot a(n) + 0 \cdot a(2n + 1)
\]

Then

\[
M_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\]
Corollaries...

- In a sense, constant-recursive sequences are “1-regular”.

- A $k$-regular sequence has constant-recursive subsequences. For example, $a(k^n) = \lambda M_1 M_0^n \kappa$.

- If $a(n)$ is $k$-regular, then $a(n) = O(n^d)$ for some $d$. 
A sequence $a(n)$ is *$k$-automatic* if there is a finite automaton whose output is $a(n)$ when fed the base-$k$ digits of $n$.

The Thue–Morse sequence is 2-automatic:

![Automaton Diagram](image-url)

Allouche–Shallit 1992:
A $k$-regular sequence is finite-valued if and only if it is $k$-automatic.
Fix $k \geq 2$. 
Automatic sequences have been very well studied.

Charlier–Rampersad–Shallit 2011: Many operations on $k$-automatic sequences produce $k$-regular sequences.

Büchi 1960: If $a(n)$ is eventually periodic, then $a(n)$ is $k$-automatic for every $k \geq 2$. 
Hierarchy of integer sequences

We add eventually periodic sequences:

\[ k\text{-regular} \rightarrow k\text{-automatic} \rightarrow \text{ev. periodic} \]
The sequence $a(n) = n$ is $k$-regular for every $k \geq 2$:

$$a(kn + i) = k(1 - i)a(n) + i a(kn + 1)$$
$$a(k^2n + i) = k(k - i)a(n) + i a(kn + 1)$$

It follows that every polynomial sequence is $k$-regular (as every polynomial sequence is constant-recursive).
Every (eventually) polynomial sequence is $k$-regular.
If $a(n)$ is eventually polynomial and $k$-automatic, then $a(n)$ is eventually constant.
Every polynomial sequence is constant-recursive. (And not every $k$-automatic sequence is constant-recursive.)
Allouche–Shallit 1992:
If \( a(n) \) is constant-recursive and \( k \)-regular, then \( a(n) \) is eventually quasi-polynomial.

Hierarchy of integer sequences:

- \( \text{ev. constant-recursive} \)
- \( \text{ev. polynomial} \)
- \( \text{\( k \)-automatic} \)
- \( \text{ev. quasi-polynomial} \)
Hierarchy of integer sequences

And to entice us...
Sequences that are not regular

- By Bell’s theorem, \( \nu_2(n^2 + 7) \) is not 2-regular.

  \[ 0 \ 3 \ 0 \ 4 \ 0 \ 5 \ 0 \ 3 \ 0 \ 3 \ 0 \ 7 \ 0 \ 4 \ 0 \ 3 \ 0 \ 3 \ 0 \ 4 \ 0 \ 6 \ 0 \ 3 \ 0 \ 3 \ 0 \ 5 \ 0 \ 4 \ 0 \ 3 \ \cdots \]

- Bell \( \sim 2005 \), Moshe 2008, Rowland 2010:

  \( \lfloor \alpha + \log_k(n + 1) \rfloor \) is \( k \)-regular if and only if \( k^\alpha \) is rational.

  For example, \( \lfloor \frac{1}{2} + \log_2(n + 1) \rfloor \) is not 2-regular.

  \[ 0 \ 1 \ 2 \ 2 \ 2 \ 3 \ 3 \ 3 \ 3 \ 3 \ 4 \ 4 \ 4 \ 4 \ 4 \ 4 \ 4 \ 4 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ 5 \ \cdots \]

Is there a natural (larger) class that these sequences belong to?
Two generalizations of $k$-regular sequences:

- Allow polynomial coefficients in $n$ (analogous to polynomial-recursive sequences).
- Becker 1994, Dumas 1993, Randé 1992:
  If $a(n)$ is $k$-regular, then $f(x) = \sum_{n=0}^{\infty} a(n)x^n$ satisfies a Mahler functional equation
  \[ \sum_{i=0}^{m} p_i(x)f(x^{ki}) = 0. \]

How natural are these generalizations?

It remains to be seen. . .