A sequence is k-*automatic* if its n'th term is generated by a finite state machine with n in base k as the input.

Examples of automatic sequences

The Thue-Morse sequence

$$01101001100101101001 0 \cdots$$

This sequence is 2-automatic.

e.g., n=13, then $n = 1101$ in base 2. Output $= 1$.

The Rudin-Shapiro sequence.

$$1, 1, 1, -1, 1, 1, -1, 1, 1, 1, 1, -1, -1, \ldots.$$

The n'th term of this sequence is given by -1 to the number of occurrences of "11" in the binary exmpansion of n.

This sequence is 2-automatic.

A periodic sequence is k-automatic for every k. Also, a sequence which is eventually periodic is k-automatic for every k.

Integers $p, q > 1$ are called *multiplicatively dependent* if $p^a = q^b$ for some positive integers a, b; otherwise, they are called multiplicatively independent.

THEOREM: (COBHAM) If a sequence is both p-automatic and q-automatic and p and q are multiplicatively independent, then the sequence is eventually periodic.

The *kernel* of a k-automatic sequence,

$$\{f(n) \mid n \geq 0\},$$

(or k-kernel) is the set of all subsequences of the form

$$\{f(k^a n + b) \mid n \geq 0\}$$

with $a \geq 0$ and $0 \leq b < k^a$.

Example: Take the Thue-Morse sequence

$$01101001100101101001 0 \cdots .$$

If $TM(n)$ denotes the n'th term of the sequence, then

$$TM(2n) = TM(n) \qquad TM(2n+1) = 1 - TM(n).$$

Thus either

$$TM(2^a n + b) = TM(n)$$

for all n, or

$$TM(2^a n + b) = 1 - TM(n)$$

for all n.

The 2-kernel of the Thue-Morse sequence consists of only two sequences; namely, the Thue-Morse sequence and its "opposite."

In general, we expect the k-kernel of a sequence to be infinite.

In the case of k-automatic sequences, however, the k-kernel is finite.

THEOREM: A sequence is k-automatic if and only if its k-kernel is finite.

Allouche and Shallit used the k-kernel characterization of automatic sequences to give a natural generalization: Regular sequences.

Notice that the collection of (real) sequences forms a $\mathbb{Z}$-module, which we call S.

$$(1, 3, \pi, 2, 0, \ldots) + (0, 1, 1, -1, 2, \ldots)$$

$$= (1, 4, \pi + 1, 1, 2, \ldots).$$

$$3 \cdot (1, 3, \pi, 2, 0, \ldots) = (3, 9, 3\pi, 6, 0, \ldots).$$

Given a real sequence $\{f(n) \mid n \geq 0\}$, we let $M(f, k)$ denote the $\mathbb{Z}$-submodule of S generated by all sequences in the k-kernel of $\{f(n)\}$.

Definition: We say that a sequence $\{f(n)\}$ is k-*regular* if the the module $M(f,k)$ is finitely generated.

Remark: A k-automatic sequence is k-regular.

After all, if the k-kernel is finite, then clearly the module $M(f,k)$ is finitely generated.

Some examples of regular sequences

Let $p(x)$ be a polynomial with real coefficients. Then the sequence $p(0), p(1), \ldots$ is k-regular for every k.

More generally, if

$$a_0 + a_1 x + a_2 x^2 + \cdots$$

is a rational power series with no poles inside the unit disc, then $a_0, a_1, \ldots$ is k-regular for every k.

An example from history?

- In the first century, Josephus along with 40 other rebels were hiding in a cave from the Romans during the Roman-Jewish war.

- Faced with certain death, the 41 men decided killing themselves was preferable to being killed by the Romans.

- Suicide was considered much worse than murder in Judaism (according to a book I read about this, anyway)

- The men decided to form a circle and kill every other person in the circle till only one was left, the last person would then commit suicide.

Let $J(n)$ denote the last person to die in the Josephus circle of size n.

Then one sees that

$$J(2n) \;=\; 2J(n) - 1$$

$$J(2n+1) \;=\; 2J(n) + 1.$$

These relations show that the sequence $\{J(n)\}$ is 2-regular.

Incidentally, when $n = 41$,

$$J(41) = 2J(20) + 1 = 2(2J(10) - 1) + 1 = 19.$$

A different example.

Let $f(n)$ count the number of 1's in the binary expansion of n. Then $f(n)$ is 2-regular.

To see this, notice that $f(2^a n + b)$ is just the number of 1's in the binary expansion of b added to the number of 1's in the binary expansion of n. Thus $f(2^a n + b) = f(n) + f(b)$.

It follows that $M(f, 2)$ is generated by

$$(f(0), f(1), \ldots)$$

and the constant sequence

$$(1, 1, 1, \ldots).$$

As we have seen, a k-regular sequence may be unbounded, while a k-automatic sequence only takes on finitely many values.

THEOREM: A k-regular sequence is k-automatic if and only if it only takes on finitely many values.

CLOSURE PROPERTIES:

- The sum of k-regular sequences is k-regular.

- The coordinate-wise product of two k-regular sequences in k-regular

- The Cauchy product of two k-regular sequences is k-regular.

- If $\{f(n)\}$ is k-regular, so is $\{f(an+b)\}$.

NOTE: If $(a_i)_{i\geq 0}$ and $(b_i)_{i\geq 0}$ are two sequences, then $c_i := a_i b_i$ is their coordinate-wise product and $C_i := a_i b_0 + a_{i-1} b_1 + \cdots + a_1 b_{i-1} + a_0 b_i$ is their Cauchy product.

The Allouche-Shallit project: To understand regular sequences.

Picture:

Regular seq.

Automatic seq.

eventually periodic

rational seq. of poly. growth

QUESTIONS:

Are there analogues of well-known theorems about rational functions, automatic sequences for regular sequences?

Example: Is there an analogue of Cobham's theorem?

Yes, there is an analogue of Cobham's theorem.

The correct analogue is that a sequence which is both k-regular and ℓ-regular is rational if k and ℓ are multiplicatively independent.

Cobham's theorem is difficult to prove; fortunately, one can reduce this generalization to the original and hence rely on his theorem.

LOCAL-GLOBAL PRINCIPLE?

THEOREM: Let $\{f(n)\}$ be a k-regular sequence taking integer values and let $\{g(n)\}$ be an integer sequence. Suppose that for infinitely many prime numbers p there is a sequence $h_p(n) \in M(f,k)$ such that $g(n) \equiv h_p(n) \bmod p$ for all n. Then there is an integer m such that $mg(n) \in M(f,k)$.

PROOF: Let $f_1, \ldots, f_d$ be a minimal set of generators for $M(f,k)$ and consider

$$\begin{pmatrix} f_1(0) & f_1(1) & f_1(2) & \cdots \\ f_2(0) & f_2(1) & f_2(2) & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ f_d(0) & f_d(1) & f_d(2) & \cdots \\ g(0) & g(1) & g(2) & \cdots \end{pmatrix}$$

Why is this theorem useful?

Idea is to take $g(n) = f(n+i)$. If we can use this to show that for every i, some multiple of $f(n+i)$ is in $M(f,k)$, then the finite generation will give a recurrence satisfied by $\{f(n)\}$, which will imply that it is a rational sequence.

REMARK Notice that if $\{f(n)\}$ is k-regular and ℓ-regular, then $\{f(n) \bmod p\}$ is k-automatic and ℓ-automatic since it only takes on finitely many values. Thus if k and ℓ are multiplicatively independent, then the reduced sequence is eventually periodic by Cobham's theorem.

These facts allow us to deduce

THEOREM: An integer sequence which is both k and ℓ regular is rational if k and ℓ are multiplicatively independent.

This theorem allows us to recognize k,ℓ-regular sequences, but it doesn't help with sequences which are just k-automatic.

Allouche and Shallit asked the innocent looking question:

Let $\{f(n)\}$ be an unbounded k-regular integer sequence. Must it assume infinitely many composite numbers as values?

The answer to this questions is ‘yes’. The methods used to answer this question show that rational power series and regular sequences are intimately connected.

We have seen that regular sequences can be thought of in terms of kernels and modules. There is another convenient way of thinking of regular sequences.

Let $A_0, \ldots, A_{k-1}$ be $d \times d$ integer matrices and let $\mathbf{v}(0)$ be a vector satisfying $A_0\mathbf{v}(0) = \mathbf{v}(0)$.

We define a sequence of vectors $\mathbf{v}(0), \mathbf{v}(1), \ldots$ recursively, by declaring that for $0 \leq i < k$,

$$\mathbf{v}(kn + i) = A_i\mathbf{v}(n).$$

Then the sequence of first coordinates of $\mathbf{v}(n)$ is k-regular. Moreover, any k-regular sequence can be created this way.

Example. Let $k = 2$, Let

$$\mathbf{v}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We let

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Then it is easy to understand $\mathbf{v}(n)$. Notice that

$$\mathbf{v}(2n) = A_0\mathbf{v}(n) = \mathbf{v}(n),$$

and

$$\mathbf{v}(2n+1) = A_1\mathbf{v}(n) = \mathbf{v}(n) + [1,0]^{\mathsf{T}}.$$

Thus the first coordinate of $\mathbf{v}(n)$ is just the number of 1's in its binary expansion.

So the terms of a k-regular sequence can be thought of as elements of the form

$$\mathbf{e_1}^{\mathsf{T}} X \mathbf{v}(0),$$

where X is some word on $A_0, \ldots, A_{k-1}$.

Given a finite set S of matrices, we call the collection of all products of elements of S, the semigroup on S.

The question becomes: What kind of sequences arise in this manner? That is, what are the sequences of the form $\mathbf{w}^{\mathsf{T}} X \mathbf{v}$, where X runs over the elements in a semigroup of matrices?

Special case: what happens if we have just one matrix, A.

Then the semigroup is just powers of A and we are looking at sequences of the form

$$\mathbf{w}^{\mathsf{T}} A^n \mathbf{v}.$$

THEOREM: (Schützenberger) The power series

$$\sum_{n=0}^{\infty} \mathbf{w}^{\mathsf{T}} A^n \mathbf{v} t^n$$

is a rational power series.

THEOREM: Let S be a semigroup of matrices and let $\mathbf{v}, \mathbf{w}$ be vectors. Then either

$$\{\mathbf{w}^\top A\mathbf{v} \mid A \in S\}$$

is finite or it contains all the members of an unbounded rational sequence.

In particular, if we look at an infinite subset of the integers and look at regular sequences whose values are restricted to this set, then if our set does not contain all the members of an unbounded rational sequence, then our regular sequence **MUST** be bounded.

Recall that a bounded regular sequence of integers is automatic. Thus we can restate this remark as

THEOREM: Let T be a set of integers which does not contain an unbounded rational sequence. Then any regular sequence whose values are confined to T is necessarily automatic.

Examples. If T grows super-exponentially or T is the set of primes, then T does not contain an unbounded rational sequence.

REMARK: Any rational sequence whose values are all prime numbers is necessarily bounded.

Why? Use Schützenberger's result. Then

$$\mathbf{e_1}^{\mathsf{T}} A^n \mathbf{v} = \text{prime} \qquad \forall n.$$

WOLOG A is invertible. Look at the case $\det(A) = 1$.

A curious dichotomy

Using a theorem of Burnside about semigroups of matrices, it can be shown that the following dichotomy holds.

If $f(n)$ is a k-regular sequence, then either

$$f(n) \;=\; O((\log n)^d)$$

or there is some $\alpha > 0$ such that

$$f(n) > n^\alpha \qquad \text{for infinitely many } n.$$

For instance, it is impossible to have

$$f(n) \;\sim\; e^{\sqrt{\log\, n}}.$$

In light of this dichotomy, we say that if a k-regular sequence $f(n)$ satisfies

$$f(n) = \mathsf{O}((\log\ n)^d)),$$

then $f(n)$ is a *polylog bounded* (p.l.b.) k-regular sequence.

Notice that polylog bounded regular sequences contain automatic sequences. It seems reasonable that one can give a concrete description of p.l.b. regular sequences in terms of automatic sequences.

Examples of p.l.b. regular sequences

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

is a p.l.b. 2-regular sequence.

Notice that if $f(n)$ denotes the n'th term of this sequence, then $f(n) = a$ whenever $2^a \leq n < 2^{a+1}$.

Hence $f(2^i n + j) = f(n) + i$ for $0 \leq j < 2^i$. Consequently $M(f; 2)$ is spanned by $f(n)$ and the constant sequence 1 as a Z-module. Also, $f(n) \leq \log_2 n$ and so it is p.l.b.

The earlier example, in which $f(n)$ is the number of 1's in the binary expansion of n is also a p.l.b. 2-regular sequence.

The p.l.b. sequences have the following property:

P.L.B. SEQUENCES

$\cap$

RATIONAL SEQUENCES

$=$

EVENTUALLY PERIODIC SEQUENCES

As with regular sequences, p.l.b. regular sequences have the following closure properties.

- They are closed under coordinate-wise sum.
- They are closed under coordinate-wise product.
- They are **NOT** closed under Cauchy product.

The $\star$ product

In an attempt to describe p.l.b. regular sequences, we define an associative, noncommutative product on regular sequences.

Let $f(n)$ and $g(n)$ be two k-regular sequences. Given $n \geq 1$ we let $a_0 a_1 \cdots a_d$ denote its base k representation.

$$(f \star g)(a_0 \cdots a_d)$$

$$= \sum_{k=0}^{d} f(a_0 \cdots a_k) g(a_0 a_{k+1} \cdots a_d).$$

Example.

Let $k = 2$ and let $n = 19$. Then $[n]_2 = 10011$.

Define

$$(f \star g)(10011) = f(10011)g(1) + f(1001)g(11)$$

$$+f(100)g(111)+f(10)g(1011)+f(1)g(10011).$$

In base 10, we have

$$(f \star g)(19) = f(19)g(1) + f(9)g(3)$$

$$+f(4)g(7) + f(2)g(11) + f(1)g(19).$$

Example. Let $k = 2$. Let $f(n) = g(n) = 1$ for all n.

Then $(f \star g)(n)$ is just the number of bits in n.

Let $k = 2$. $f(n) = 1$ if n is odd and 0 if n is even. Let $g(n) = 1$ for all n. Then

$$(f \star g)(n)$$

is just the number of 1's in the base 2-representation of n.

THEOREM If $f(n)$ and $g(n)$ are k-regular sequences, then so is $f \star g$; moreover, if f and g are p.l.b., then so is $f \star g$.

Let A denote the collection of k-regular sequences. Then

$$(A, \star, +)$$

is a noncommutative, associative algebra.

$$(B, \star, +)$$

is a subalgebra, where B is all p.l.b. k-regular sequences.

THEOREM: B is generated as a $\mathbb{Q}$-algebra by the collection of automatic sequences.

This means that any p.l.b. k-regular sequence is a $\mathbb{Q}$-linear combination of things of the form $f_1 \star f_2 \star \cdots \star f_d$ with $f_1, \ldots, f_d$ automatic, $d \geq 1$.

Thus we have a structure theorem for p.l.b. regular sequences.

Can we use this structure theorem to answer the questions of Allouche and Shallit about regular sequences in the p.l.b. case?

ANSWER: Perhaps.

Remarks about the algebra $(A, \star, +)$

- it has a multiplicative identity. $\mathsf{id}(n) = 1$ if $1 \leq n < k$, and $\mathsf{id}(n) = 0$ otherwise.

- It is not a domain. For example, let $k = 3$ and let $f(n) = 1$ if n is a power of 3 and let $f(n) = 0$ otherwise. Let $g(n) = 1$ if $n = 2 \cdot 3^a$ and 0 otherwise. Then $f \star g = 0$.

- A has no nonzero nilpotent elements. This implies every prime homomorphic image of A is a domain. (For noncommutative rings R, we say that R is prime if $xRy = 0$ for some $x, y \in R$ then either $x = 0$ or $y = 0$.

- A embeds into a direct product of (not necessarily commutative) domains.

Questions about the algebras A and B.

Does A or B contain a copy of the free algebra on two generators?

This question has to do with Goldie's work on noncommutative localization. If we knew that B did not contain a copy of the free algebra, then we could embed B in a semisimple Artinian ring. We may then think of automatic sequences in terms of matrices over some division ring. This might be useful.