THE MORSE SEQUENCE

RICHARD G. SWAN

ABSTRACT. We give a proof that the Morse sequence has no 3 times repeated block and derive from it a sequence on 3 letters that has no repeated block.

1. THE MORSE SEQUENCE

The Morse sequence is a sequence of a's and b's defined as follows: We start with the sequence $S_0 = \{a\}$.If S_n has been defined having 2^n elements we define S_{n+1} to be $S_n\overline{S_n}$ where, if S is a sequence of a's and b's, $\overline{S}$ denotes the sequence obtained from S by interchanging the a's and b's. The sequence is the union of these blocks and thus looks like

$$abbabaabbaababba\ldots$$

or, using vertical bars to separate the blocks S_n as in [1],

$$a \mid b \mid ba \mid baab \mid baababba \mid \ \ldots$$

.

In [1] Jacobson makes use of the following interesting property to construct an infinite semigroup with zero generated by two elements in which $x^3 = 0$ for all elements x.

Theorem 1.1. *There is no non–void block U of a's and b's such that UUU occurs in the Morse sequence.*

For the proof, Jacobson refers to the paper [2] of Morse and Hedlund, but after considerable searching I was unable to find this theorem in that paper. I therefore worked out my own version of the proof which I will present here for whatever it is worth. In the course of this I noticed that one can define a sequence on 3 letters which has no repeated block UU. Starting with the Morse sequence delete the first letter a and separate the remainder of the sequence into pairs:

$$bb\ ab\ aa\ bb\ aa\ ba\ bb\ \ \ldots$$

Then replace each pair aa by x, each bb by y, and replace each ab and each ba by z getting a sequence

$$y\ z\ x\ y\ x\ z\ y\ \ \ldots$$

.

Theorem 1.2. *There is no non–void block U of x's, y's and z's such that UU occurs in the sequence just defined.*

As in [1] it follows that there is an infinite semigroup with zero generated by three elements in which $x^2 = 0$ for all elements x. There is no such sequence or semigroup on two generators since if no aa or bb occurs the sequence must consist of alternate a's and b's.

2. Inventory

We will need some preliminary observations on short blocks contained in the Morse sequence. I think the clearest way to present these is just to make a list of all the possible blocks. The results will be clear from this.

We denote the terms of the Morse sequence by $f(0)f(1)f(2)\dots$. The definition of the Morse sequence can be expressed by $f(0) = a$, and, for $2^i \leq n < 2^{i+1}$, $f(n) = \overline{f}(n - 2^i)$ where $\overline{a} = b$ and $\overline{b} = a$ as above. A block of the Morse sequence will always mean a segment U of the form $f(i)f(i+1)\dots f(j)$. We write $|U|$ for the length $j - i + 1$ of U. An n-block will mean one of length n. By the parity of an element f(i) of the Morse sequence I will mean $i \mod 2$ and the parity of a block $U = f(i)f(i+1)\dots$ will mean that of the first element $f(i)$.

Lemma 2.1. *A block of the Morse sequence of the form $f(4i)f(4i+1)f(4i+2)f(4i+3)$ is either abba or baab.*

Proof. This is clear for the block S_2. Suppose that the given block lies in $S_{n+1} = S_n\overline{S_n}$ for $n \geq 2$. The first element of $\overline{S_n}$ is $f(2^n)$ with $4|2^n$ so the given block lies in either S_n or $\overline{S_n}$ and we use induction on n. □

It follows that an 8-block of the form $f(4i)f(4i+1)\dots f(4i+7)$ must have one of the forms

$$\underline{a}b\underline{b}a\ \underline{a}b\underline{b}a \quad \underline{a}b\underline{b}a\ \underline{b}a\underline{a}b \quad \underline{b}a\underline{a}b\ \underline{a}b\underline{b}a \quad \underline{b}a\underline{a}b\ \underline{b}a\underline{a}b$$

where the elements of even parity are underlined. Since any block of length at most 4 lies in such an 8-block we obtain the following inventory of small blocks.

Length 2, even parity: *ab*, *ba*.
Length 2, odd parity: *ab*, *ba*, *aa*, *bb*.
Length 3, even parity: *abb*, *baa*, *aba*, *bab*
Length 3, odd parity: *aab*, *bba*, *aba*, *bab*
Length 4, even parity: *abba*, *baab*, *baba*, *abab*
Length 4, odd parity: *aabb*, *bbaa*, *aaba*, *bbab*, *abaa*, *babb*

From this it is easy to check the following facts.

Lemma 2.2. (1) *A block of length 2 and even parity has the form ab or ba.*
(2) *The Morse sequence has no block of the form aaa or bbb.*
(3) *Let U be a block of length 3 not aba or bab. Then all occurrences of U in the Morse sequence have the same parity.*
(4) *If U is a block with $|U| \geq 4$ then all occurrences of U in the Morse sequence have the same parity.*

For the last statement it suffices to look at the first 4 letters in U.

Corollary 2.3. *If $|U| \geq 3$ is odd and U is not aba or bab then no UU occurs in the Morse sequence.*

This follows from the fact that the two occurrences of U would have different parities.

3. Proof of Theorem 1.1

We will prove the following stronger result.

Theorem 3.1. *Let x be the first letter of $U = x\ldots$. Then UUx does not occur in the Morse sequence.*

We first dispose of the case in which $|U|$ is odd. This follows from Corollary 2.3 unless $U = aba$ or $U = bab$. If $U = aba$, then $x = a$ and $UUx = abaabaa$ which is impossible by Corollary 2.3 applied to baa. A similar argument applies if $U = bab$.

We now prove the theorem by induction on $|U|$. The case $|U| = 1$ is clear. Define two subsequences of the Morse sequence by $\mathcal{M}_{\text{even}} = f(0)f(2)f(4)\ldots$ and $\mathcal{M}_{\text{odd}} = f(1)f(3)f(5)\ldots$. The sequence $\mathcal{M}_{\text{even}}$ looks exactly like the original Morse sequence while $\mathcal{M}_{\text{odd}}$ looks like the original Morse sequence with the a's and b's interchanged. These observations follow from the following lemma.

Lemma 3.2. *$f(2n) = f(n)$ and $f(2n+1) = \overline{f}(n)$.*

Proof. This is clear for $n = 0$. We use induction on n. Suppose the result holds for $n < 2^i$ and suppose that $2^i \le n < 2^{i+1}$. Then $2^{i+1} \le 2n < 2^{i+2}$ so $f(2n) = \overline{f}(2n - 2^{i+1}) = \overline{f}(2(n - 2^i))$. Since $n - 2^i < 2^i$, the induction hypothesis shows that this is $\overline{f}(n - 2^i) = f(n)$. Since $f(2n+1) = \overline{f}(2n)$ by Lemma 2.2(1) the second statement follows. □

Suppose now that UUx occurs with $|U|$ even and with even parity. Let $V = U \cap \mathcal{M}_{\text{even}}$. Then $V = x\ldots$ and $\mathcal{M}_{\text{even}}$ contains the block $UUx \cap \mathcal{M}_{\text{even}} = VVx$. Since $|V| = |U|/2 < |U|$, this is impossible by the induction hypothesis. If UUx occurs with $|U|$ even and with odd parity we use the same argument on $V = U \cap \mathcal{M}_{\text{odd}}$.

4. Related sequences

In order to construct sequences with no repeated block we make use of the following consequence of Theorem 3.1.

Corollary 4.1. *If the non–void block U has even length, the Morse sequence contains no block UU of odd parity.*

Proof. Let $U = xVy$ and let z be the letter following the second U in the block UU so the Morse sequence contains the block $xVyxVyz$. The x's have odd parity and the y's have even parity. By Lemma 2.2(1), $x = \overline{y} = z$ so the Morse sequence contains the block UUx contradicting Theorem 3.1. □

We define a new sequence on 4 elements as follows. Omit the first term $f(0)$ of the Morse sequence and divide the remaining elements into pairs $f(1)f(2) \mid f(3)f(4) \mid \ldots$. Replace each pair $f(2n-1)f(2n)$ by a single letter as follows. Replace aa by x, bb by y, ab by u, and ba by v, getting a sequence $yuxyxvy\ldots$.

Corollary 4.2. *This sequence contains no non–void block of the form UU.*

Proof. By replacing x by aa, y by bb, u by ab, and v by ba, we get a block of the Morse sequence of the form VV of odd parity and with $|V|$ even contradicting Corollary 4.1. □

The following observation enables us to produce a sequence on 3 elements with the same property.

Lemma 4.3. *In the sequence just defined u always occurs in a block yux and v always occurs in a block xvy.*

Proof. Each u is the image of a block ab of the Morse sequence of odd parity. By Lemma 2.2(1), the letter before the a must be b. By Corollary 4.2 the letter before that b must be b otherwise we would get a block uu. Similarly the letter after our ab must be a and the letter after that must also be a. A similar argument applies to v. □

The sequence of Theorem 1.2 is obtained from the present sequence by replacing all u's and v's by z's. By Lemma 4.3 no zz can occur and by Corollary 4.2 no xx or yy can occur so there is no block UU with $|U| = 1$. Any longer block lifts uniquely to a block of our 4-element sequence since xz must lift to xv, yz to yu, zx to ux and zy to vy. Each of the two blocks U will lift to a block V of the 4 element sequence giving us a block VV which contradicts Corollary 4.2. This proves Theorem 1.2.

To conclude I will point out an alternative construction for this sequence: We start with the Morse sequence replacing a by y and b by x. Call this sequence $\mathcal{M}_0$. We then insert a letter between each consecutive pair of letters of $\mathcal{M}_0$ according to the following rules: Between two consecutive x's of $\mathcal{M}_0$ we insert a y, between two consecutive y's of $\mathcal{M}_0$ we insert an x, and between each xy or yx of $\mathcal{M}_0$ we insert a z.

To see that this gives the same sequence let $g(0)g(1)g(2)\dots$ be the sequence of Theorem 1.2. Then $g(n)$ is determined by $f(2n+1)f(2n+2)$ so the elements $g(2n)g(2n+1)g(2n+2)$ are determined by the block

$$f(4n+1)f(4n+2)f(4n+3)f(4n+4)f(4n+5)f(4n+6)$$

. Using Lemma 3.2 this block is easily seen to be

$$\overline{f}(n)\overline{f}(n)f(n)f(n+1)\overline{f}(n+1)\overline{f}(n+1)$$

so the elements $g(2n)g(2n+1)g(2n+2)$ are determined by $f(n)$ and $f(n+1)$ according to the table

f(n)	f(n+1)	g(2n)	g(2n+1)	g(2n+2)
a	a	y	x	y
a	b	y	z	x
b	a	x	z	y
b	b	x	y	x

It follows that $g(0)g(2)g(4)\dots$ is the Morse sequence with y and x in place of a and b and $g(2n+1)$ is determined by the rules given above.

References

1. N. Jacobson, Structure of Rings, Amer. Math. Soc. Colloquium Publications XXXVII, Providence RI 1956.
2. M. Morse and G. A. Hedlund, Symbolic dynamics, Amer. J. Math. 60(1936), 815–866.

Department of Mathematics, The University of Chicago, Chicago, IL 60637
E-mail address: swan@math.uchicago.edu