MATH CIRCLES 2016: PATTERS IN WORDS 3

BLAKE MADILL

Abstract. This is lecture three of a three part lecture series on infinite sequences and patterns in words at the University of Waterloo. It will take place on March 9, 2016.

1. Recap from last time!

Welcome back! Let's recap what we did last time.

To understand an infinite word, it is quite often sufficient to understand its finite subwords. For example, to see that the Thue-Morse word contains a square you can just consider all the length-2 subwords of t: $\{01, 10, 11, 00\}$. Namely, 11 is a square in t. Let us start with a basic question about subwords:

How many subwords of length n does a given infinite word have?

We make the following definition.

Definition 1. *Let w be an infinite word over the alphabet Σ. For $n \in \mathbb{N} \cup \{0\}$ we denote the number of subwords of w of length n by $\rho_w(n)$. We call the ρ_w the subword complexity function of w.*

Theorem 1. *Let w be an infinite word on the alphabet Σ. Then $\rho_w(n) \leq |\Sigma|^n$ for every $n \in \mathbb{N}$.*

Theorem 2. *Let w be an infinite word on Σ. Then $\rho_w(n) \leq \rho_w(n+1) \leq |\Sigma|\rho_w(n)$.*

2. Periodic Words

In this section we discuss the subword complexity of periodic words. Recall that we usually says a mathematical object is periodic if some sort of repetition continues to happen forever. For example, you likely know that the function $\sin(x)$ is periodic with period 2π. We thus form the following natural definition.

Definition 2. *Let z be a nonempty finite word. We denote by z^ω, the infinite word $z^\omega = zzzzz\cdots$. We call such a word* periodic*. We say that an infinite word w is* ultimately periodic *if $w = uz^\omega$, where u is a possibly empty finite word and z is a non-empty finite word. Clearly, any periodic word is ultimately periodic.*

Example 1. *If*

$$w_1 = 01100110011001100110\cdots$$

and

$$w_2 = 11011001100110011 0\cdots$$

then we see that $w_1 = (0110)^\omega$ is periodic and $w_2 = 11(0110)^\omega$ is ultimately periodic.

Definition 3. *Let w be a finite word. We denote the length of w (i.e. the number of letters in w) by $\ell(w)$.*

What is really going on in the above exercise is the following theorem.

Theorem 3. *Let $w = uz^\omega$ be an ultimately periodic word. Assume z cannot be written as a power of a shorter word. Moreover, assume that u is either empty or ends with a letter different from the last letter of z. (Any ultimately periodic word can be written in this way). Then $\rho_w(n) = \ell(uz)$, for all $n \geq \ell(uz)$. That is, the sequence $(\rho_w(n))_{n=1}^\infty$ is eventually constant.*

From the above theorem, we see that the subword complexity function of an ultimately periodic word w is bounded. That is, $\rho_w(n)$ does NOT continue to get bigger and bigger. Here is a natural question:

If w is an infinite word such that $\rho_w(n)$ is bounded, must w be ultimately periodic?

As it turns out, the answer to this question is "yes". We say a sequence $(a_n)_{n=1}^\infty$ of real numbers is increasing if $a_n \leq a_m$ whenever $n \leq m$. We say that this sequence is strictly increasing if $a_n < a_m$ whenever $n < m$. From Theorem 4 we know that $(\rho_w(n))_{n=1}^\infty$ is increasing. In fact, it is strictly increasing.

Theorem 4. *(Morse and Hedlund) Let w be an infinite word. Then either w is ultimately periodic or $(\rho_w(n))_{n=1}^\infty$ is strictly increasing.*

Corollary 1. *Let w be an infinite word. If there exists an n such that $\rho_w(n) = \rho_w(n+1)$ then w is ultimately periodic.*

Exercise 1. *For each of the functions $f(n)$, find an infinite word w such that $\rho_w(n) = f(n)$ for every $n \in \mathbb{N}$, if possible. If no such example exists, explain why not.*

(1) $f(n) = 3$
(2) $f(n) = n$
(3) $f(n) = n^2 - n - 1$
(4) $f(n) = 3^n$
(5) $f(n) = n + 1$

As you probably just saw, (1) through (4) over the above problem weren't too hard. However, you likely had trouble constructing an example of an infinite word w such that $\rho_w(n) = n+1$ for every $n \in \mathbb{N}$. Such words are called *Sturmian words*. These are words which have the smallest possible subword complexity without being ultimately periodic. We break up this problem into exercises.

Exercise 2. *If w is Sturmian, how many letters occur in w?*

Let $f_0 = 0$ and let $f_1 = 01$. Define, for $n \geq 2$, $f_n = f_{n-1}f_{n-2}$. We define an infinite word f over the alphabet $\{0,1\}$ by requiring that each f_n is a prefix of f. Observe that

$$\begin{aligned} f_2 &= 010, \\ f_3 &= 01001, \\ f_4 &= 01001010 \\ f_5 &= 0100101001001, \end{aligned}$$

so that $f = 0100101001001\cdots$.

Exercise 3. *How does this word relate to the Fibonacci sequence*

$$0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots?$$

For this reason, we call f the Fibonacci word. An alternative construction of f is given as follows. Let φ be a function defined on $\{0,1\}$ by $\varphi(0) = 01$ and $\varphi(1) = 0$. Moreover, for $w = x_1x_2\cdots x_n$, where $x_i \in \{0,1\}$, define $\varphi(w) := \varphi(x_1)\varphi(x_2)\cdots\varphi(x_n)$. Then,

$$\varphi^n(0) = f_n,$$

for $n \in \mathbb{N}$ and so f may be constructed by repeatedly applying φ to 0. If we then extend the definition of φ to infinite words (letter-by-letter) we see that $\varphi(f) = f$. In fact, we have that $\varphi(f_n) = f_{n+1}$ for every $n \in \mathbb{N} \cup \{0\}$.

Lemma 1. *Let f be the Fibonacci word and let φ be as above. If v is a subword of f then $\varphi(v)$ is a subword of f.*

Lemma 2. *Let f be the Fibonacci word. If v is a subword of f then there exists a subword of f, u, such that v is a subword of $\varphi(u)$.*

We are now ready to begin our quest to prove the following theorem.

Theorem 5. *The Fibonacci word is Sturmian.*

Exercise 4. *For a finite word* $u = x_1 x_2 \cdots x_n$, $x_i \in \{0, 1\}$, *define* $u^R = x_n \cdots x_2 x_1$. *We call* u^R *the* reversal *of* u. *Show that if* u *is a subword of* f *then* u^R *is a subword of* f.

Exercise 5. *Show that if* $\varphi(w)0$ *is a subword of* f *then* w *is a subword of* f.

Exercise 6. *We say that a finite word u is a left-special-subword of f if u is a subword of f such that $0u$ and $1u$ are subwords of f. We say that u is a right-special-subword of f if u is a subword of f such that $u0$ and $u1$ are both subwords of f. Show that the left-special-subwords of f are just the prefixes of f.*

Exercise 7. *Show that the right-special-subwords of* f *are just the reversals of the prefixes of* f*.*

Exercise 8. *Show that an infinite word* w *is Sturmian if and only if it has exactly one special-right-subword of each length.*

Exercise 9. *Conclude that* f *is Sturmian.*