Repetitions in Words—Part II

Narad Rampersad

Department of Mathematics and Statistics
University of Winnipeg

Deciding if a fixed point is periodic

- Recall: a primitive morphism generates either periodic words or words avoiding t-powers for some t.
- Given a morphism (not necessarily primitive), can we tell if its fixed points are periodic?
- Pansiot (1986) and also Harju and Linna (1986) gave a decision procedure.
- Honkala (2008) gave a a nice, short proof.

Elementary morphisms

- Honkala's proof uses the notion of an elementary morphism (Ehrenfeucht and Rozenberg 1978)
- A morphism $h : X^* \to Y^*$ is simplifiable if there is an alphabet Z, smaller than X, and morphisms $f : X^* \to Z^*$ and $g : Z^* \to Y^*$, such that $h = gf$.
- e.g., $h : 0 \mapsto 0012, 1 \mapsto 12, 2 \mapsto 012$ is simplifiable via $f : 0 \mapsto aab, 1 \mapsto b, 2 \mapsto ab$ and $g : a \mapsto 0, b \mapsto 12$.
- a morphism is elementary if it is not simplifiable

Properties of elementary morphisms

- elementary morphisms are injective on finite and infinite words (Ehrenfeucht and Rozenberg 1978)
- they are also non-erasing
- Honkala's method for deciding periodicity of fixed points is based on two cancellation lemmas
- the first extends a result of Ehrenfeucht and Rozenberg

First cancellation result

Cancellation Lemma A (Honkala 2008)

Let $h : X^* \to X^*$ and let $\mathbf{u}$ and $\mathbf{v}$ be infinite words over X. If there is a positive integer n such that $h^n(\mathbf{u}) = h^n(\mathbf{v})$, then $h^{|X|}(\mathbf{u}) = h^{|X|}(\mathbf{v})$.

The elementary case is immediate

- by induction on $|X|$
- the result is true when $|X| = 1$
- let $|X| = k > 1$ and suppose the result holds for all smaller alphabets
- if h is elementary then h is injective and the result follows immediately
- if h is not elementary then $h = gf$, where $f : X^* \to Y^*$, $g : Y^* \to X^*$, and $|Y| < |X|$.

Applying the induction hypothesis

- we have $(gf)^n(\mathbf{u}) = (gf)^n(\mathbf{v})$
- apply f to both sides: $(fg)^n f(\mathbf{u}) = (fg)^n f(\mathbf{v})$
- note: fg is a morphism from Y^* to Y^*
- apply the induction hypothesis to fg:
  $(fg)^{|Y|} f(\mathbf{u}) = (fg)^{|Y|} f(\mathbf{v})$
- apply g to both sides: $(gf)^{|Y|+1}(\mathbf{u}) = (gf)^{|Y|+1}(\mathbf{v})$
- i.e., $h^{|Y|+1}(\mathbf{u}) = h^{|Y|+1}(\mathbf{v})$
- apply h to both sides as many times as needed to get
  $h^{|X|}(\mathbf{u}) = h^{|X|}(\mathbf{v})$

Second cancellation result

Cancellation Lemma B (Honkala 2008)

Let $h : X^* \to X^*$ and let $\mathbf{u}$ be an infinite word over X. If there is a positive integer n such that $\mathbf{u} = h^n(\mathbf{u})$, then $\mathbf{u} = h^{|X|!}(\mathbf{u})$.

Another induction

- by induction on $|X|$
- the result is true when $|X| = 1$
- let $|X| = k > 1$ and suppose the result holds for all smaller alphabets
- let $Z = \{a \in X : |h^n(a)| = 1 \text{ for all } n\}$
- when h is elementary it induces a permutation of Z
- in particular $h^{|X|!}$ fixes any element of Z

The elementary case

- suppose h is elementary
- if $\mathbf{u}$ contains only letters from Z then $\mathbf{u} = h^{|X|!}(\mathbf{u})$ is immediate
- so write $\mathbf{u} = wb\mathbf{v}$ where $w \in Z^*$ and $b \notin Z$
- then $h^n(b\mathbf{v}) = b\mathbf{v}$
- since $b \notin Z$ and h non-erasing, this implies $(h^n)^\omega(b) = b\mathbf{v}$
- then there must be $m \leq |X|$ for which $(h^m)^\omega(b) = b\mathbf{v}$

The conclusion of the elementary case

- therefore $h^m(b\mathbf{v}) = b\mathbf{v}$
- so in particular $h^{|X|!}(b\mathbf{v}) = b\mathbf{v}$
- recall: $\mathbf{u} = wb\mathbf{v}$
- since $w \in Z^*$ we have $w = h^{|X|!}(w)$
- then $h^{|X|!}(\mathbf{u}) = h^{|X|!}(wb\mathbf{v}) = wb\mathbf{v} = \mathbf{u}$

The non-elementary case

- suppose h is not elementary and write $h = gf$, where $f : X^* \to Y^*$, $g : Y^* \to X^*$, and $|Y| < |X|$.
- we have $\mathbf{u} = (gf)^n(\mathbf{u})$
- apply f to both sides: $f(\mathbf{u}) = (fg)^n f(\mathbf{u})$
- note: fg is a morphism from Y^* to Y^*
- apply the induction hypothesis to fg: $f(\mathbf{u}) = (fg)^{|Y|!} f(\mathbf{u})$
- apply g to both sides: $(gf)(\mathbf{u}) = (gf)^{|Y|!+1}(\mathbf{u})$
- i.e., $h(\mathbf{u}) = h^{|Y|!+1}(\mathbf{u})$

The conclusion of the non-elementary case

- recall the hypothesis: $\mathbf{u} = h^n(\mathbf{u})$
- if $n < |X|$ we can easily obtain $\mathbf{u} = h^{|X|!}(\mathbf{u})$
- suppose $n \geq |X|$
- apply h^n to both sides of $h(\mathbf{u}) = h^{|Y|!+1}(\mathbf{u})$
- we get $h^{n+1}(\mathbf{u}) = h^{n+1}h^{|Y|!}(\mathbf{u})$
- apply Cancellation Lemma A to reduce the exponent: $h^n(\mathbf{u}) = h^n h^{|Y|!}(\mathbf{u})$
- substitute $\mathbf{u}$ for $h^n(\mathbf{u})$ to get $\mathbf{u} = h^{|Y|!}(\mathbf{u})$
- now it is easy to get $\mathbf{u} = h^{|X|!}(\mathbf{u})$

Some preliminaries

Let X be a k-letter alphabet. Let $h : X^* \to X^*$ be a morphism. Let $a \in X$ be a letter where $h(a) = au$ and

$$\mathbf{x} = h^\omega(a) = auh(u)h^2(u)\cdots$$

is an infinite word. Write $\mathbf{x} = w_1 w_2^q \mathbf{y}$, where

$$\begin{aligned} w_1 &= auh(u)h^2(u)\cdots h^{k-1}(u), \\ w_2^q &= h^k(u)h^{k+1}(u)\cdots h^{k+k!-1}(u), \\ \mathbf{y} &= h^{k+k!}(u)h^{k+k!+1}(u)\cdots , \end{aligned}$$

and $|w_2|$ is minimal.

Honkala's criterion

Theorem (Honkala 2008)

The word $\mathbf{x}$ is ultimately periodic if and only if there are integers $r \geq 0$, $s \geq 1$, and words w_3, w_4 satisfying

$$h(w_1) = w_1 w_2^r w_3, \quad h(w_2) = (w_4 w_3)^s, \quad w_2 = w_3 w_4.$$

Proof

- suppose $\mathbf{x}$ is ultimately periodic
- then there are $i < j$ such that
  $h^i(u)h^{i+1}(u)\cdots = h^j(u)h^{j+1}(u)\cdots$
- then $h^i(uh(u)\cdots) = h^i(h^{j-i}(u)h^{j-i+1}(u)\cdots)$
- apply Cancellation Lemma A:
  $h^k(uh(u)\cdots) = h^k(h^{j-i}(u)h^{j-i+1}(u)\cdots)$
- expand and rearrange:

$$\begin{aligned} h^k(u)h^{k+1}(u)\cdots &= h^{k+j-i}(u)h^{k+j-i+1}(u)\cdots \\ &= h^{j-i}(h^k(u)h^{k+1}(u)\cdots) \end{aligned}$$

Proof

- we have $h^k(u)h^{k+1}(u)\cdots = h^{j-i}(h^k(u)h^{k+1}(u)\cdots)$
- Apply Cancellation Lemma B:

$$\begin{aligned} h^k(u)h^{k+1}(u)\cdots &= h^{k!}(h^k(u)h^{k+1}(u)\cdots) \\ &= h^{k!+k}(u)h^{k!+k+1}(u)\cdots \end{aligned}$$

- but $w_2^q = h^k(u)h^{k+1}(u)\cdots h^{k+k!-1}(u)$ and $\mathbf{y} = h^{k+k!}(u)h^{k+k!+1}(u)\cdots$
- so $w_2^q\mathbf{y} = \mathbf{y}$
- we have $\mathbf{y} = w_2^\omega$ and thus $\mathbf{x} = w_1 w_2^\omega$

Proof

- now $h(\mathbf{x}) = \mathbf{x}$
- so $h(w_1)h(w_2)^\omega = w_1 w_2^\omega$
- this implies that $h(w_2)$ is a conjugate of a power of w_2
- i.e., $h(w_2) = (w_4 w_3)^s$ for some s, where $w_2 = w_3 w_4$
- thus $h(w_1)(w_4 w_3)^\omega = w_1 (w_3 w_4)^\omega$
- this implies that $h(w_1) = w_1 (w_3 w_4)^r w_3$ for some r
- this completes the proof

A more general result

Theorem (Durand 2011; Mitrofanov 2011)

The following problem is decidable: Given morphisms $h : X^* \to X^*$ and $\tau : X^* \to Y^*$ and a letter $a \in X$, is the word $\tau(h^\omega(a))$ ultimately periodic?

Characterizations of periodic words

- we have previously seen a characterization of periodic words:
- A word $\mathbf{x}$ is ultimately periodic if and only if there is a constant C such that for every n the number of factors of $\mathbf{x}$ of length n is at most C.
- we will give another, based on repetitions ending at each position

Fractional powers

- the exponent of a word is the ratio of its length to its minimal period
- the exponent of `toronto` is $7/5$
- we say it is a $7/5$-power
- for any rational α we can define the notion of α-power
- let $\phi = 1.618\cdots$ be the golden ratio (so $\phi^2 = 2.618\cdots$)

A characterization of periodic words

Theorem (Mignosi, Restivo, and Salemi 1998)

An infinite word $\mathbf{x}$ is ultimately periodic if and only if there is a constant N such that every prefix of $\mathbf{x}$ of length at least N ends with a repetition of exponent at least ϕ^2.

An application to perfect sets of words

- in topology, a perfect set is a closed set for which every point is a limit point
- In the context of infinite words, a set S of infinite words is perfect if every word $\mathbf{x} \in S$ has the following property: For every positive n, there is another word $\mathbf{y} \in S$ such that $\mathbf{x}$ and $\mathbf{y}$ agree on a prefix of length n.
- an equivalent reformulation: Let u be a finite word; if u has an infinite extension to a word in S, then it has two infinite extensions to words in S.

Perfect sets of repetition-free words

- several authors have studied classes of repetition-free words with respect to the property of being perfect
- notably: Shelton, Soni, and Currie
- they used the so-called fixing-block method
- a powerful, but technically difficult, method

Overlaps

- an overlap is a repetition with exponent larger than 2
- the Thue–Morse word

$$\mathbf{t} = 0110100110010110 \cdots$$

  is overlap-free (Thue 1912)
- Fife (1980) completely characterized the infinite overlap-free binary words
- an immediate consequence is that the set of infinite overlap-free binary words is perfect

Some notation

- we will show that for $\alpha \geq \phi^2 + 1$, the set of infinite α-power-free words over any alphabet of size at least two is perfect
- let $\alpha > 3$ be a real number
- let $a_1 \cdots a_k$ be a word of length k over an alphabet A
- let $\mathbf{t}_h$ denote the word obtained by deleting the first $h - 1$ symbols of the Thue–Morse word

The first extension lemma

Lemma A

If $a_1 \cdots a_k \mathbf{t}_h$ is an α-power-free extension of $a_1 \cdots a_k$, then there is another α-power-free extension of $a_1 \cdots a_k$.

Proof of the first extension lemma

- hypothesis: $a_1 \cdots a_k \mathbf{t}_h$ is α-power-free
- let v be a prefix of $\mathbf{t}_h$ of length longer than $k\alpha$
- since the set of infinite overlap-free words is perfect, there is another overlap-free word $v\mathbf{s}$
- claim: $a_1 \cdots a_k v\mathbf{s}$ is α-power-free

Proof of the first extension lemma

- suppose $a_1 \cdots a_k v\mathbf{s}$ contains an α-power z with period P
- then z starts somewhere within $a_1 \cdots a_k$ and extends into $v\mathbf{s}$
- $P < k$, since otherwise $v\mathbf{s}$ contains an $(\alpha - 1)$-power
- also z extends into $\mathbf{s}$, since otherwise z is contained in $a_1 \cdots a_k v$ which is α-power-free
- so $|z| \leq P\alpha < k\alpha$ and $|z| > |v| > k\alpha$
- we have a contradiction

The second extension lemma

Lemma B

If $a_1 \cdots a_k$ is α-power-free and does not end in an $(\alpha - 1)$-power, then there exists i such that $a_1 \cdots a_k \mathbf{t}_i$ is α-power-free.

Proof of the second extension lemma

- if $a_1 \cdots a_k \mathbf{t}_1$ is α-power-free we are done
- suppose it contains an α-power z with period P
- z begins within $a_1 \cdots a_k$ and extends into $\mathbf{t}_1$
- z begins at least $P \le k$ positions before $\mathbf{t}_1$, since otherwise $\mathbf{t}_1$ contains an $(\alpha - 1)$-power
- z extends at least P positions into $\mathbf{t}_1$, since otherwise $a_1 \cdots a_k$ ends with an $(\alpha - 1)$-power
- so the suffix of $a_1 \cdots a_k$ of length P equals the prefix of $\mathbf{t}_1$ of length P
- i.e., $a_{k-P+1} a_{k-P+2} \cdots a_k \mathbf{t}_{P+1} = \mathbf{t}_1$

Proof of the second extension lemma

- now consider the word $a_1 \cdots a_k \mathbf{t}_{P+1} = a_1 \cdots a_{k-P} \mathbf{t}_1$
- if it is α-power-free we are done
- suppose it contains an α-power z_2 with period P_2
- as before, z_2 begins at least $P_2 \le k$ positions before $\mathbf{t}_{P+1}$ and extends at least P_2 positions into $\mathbf{t}_{P+1}$
- furthermore, $P_2 > P$, since otherwise z_2 is contained in $\mathbf{t}_1$, which is overlap-free
- so the suffix of $a_1 \cdots a_k$ of length P_2 equals the prefix of $\mathbf{t}_{P+1}$ of length P_2
- i.e., $a_{k-P_2+1} a_{k-P_2+2} \cdots a_k \mathbf{t}_{P+P_2+1} = \mathbf{t}_{P+1}$

Proof of the second extension lemma

- we continue in this way and obtain a sequence of integers $P < P_2 < P_3 < \cdots$
- but each P_i is less than or equal to k
- so the sequence ends after some finite number, say q, of terms
- then for $i = P + P_2 + P_3 + \cdots + P_q$ the word $a_1 \cdots a_k \mathbf{t}_i$ is α-power-free

The set of α-power-free words is perfect

Theorem (Mignosi, Restivo, and Salemi 1998)

Let α be a real number greater than or equal to $\phi^2 + 1$. The set of α-power-free words over an alphabet of size at least 2 is perfect.

Proof of Theorem

- let $\mathbf{x} = a_1 a_2 \cdots$ be an α-power-free word
- there are infinitely many k such that $a_1 \cdots a_k$ does not end with an $(\alpha - 1)$-power, since otherwise $\mathbf{x}$ would be ultimately periodic
- for each such k, by Lemma B there is an i such that $\mathbf{y} = a_1 \cdots a_k \mathbf{t}_i$ is α-power-free
- if $\mathbf{x} \neq \mathbf{y}$ we are done
- if $\mathbf{x} = \mathbf{y}$ then Lemma A implies that $a_1 \cdots a_k$ has two α-power-free extensions
- this concludes the proof

The tree of α-power-free words

- the last result says something about infinite branches in the tree of α-power-free words
- one can also ask about finite branches in this tree
- a leaf corresponds to a maximal α-power-free word
- let t and k be positive integers
- any t-power-free word over a k-letter alphabet can be extended to a maximal t-power-free word (Bean, Ehrenfeucht, McNulty 1979)

Maximal squarefree words

Theorem (Petrova and Shur 2012)

For every n there is a ternary squarefree word w whose longest squarefree extension wv satisfies $|v| = n$.

Conclusion

- we have seen the following results:
- primitive morphisms generate either periodic words or words avoiding t-powers for some t
- there is an algorithm to decide if an infinite word generated by a morphism is periodic
- for $\alpha \geq \phi^2 + 1$, the set of α-power-free words is perfect

The End