EGMO 2012/8

Evan Chen

Twitch Solves ISL

Episode 148

Problem

A word is a finite sequence of letters from some alphabet. A word is *repetitive* if it is a concatenation of at least two identical subwords (for example, *ababab* and *abcabc* are repetitive, but *ababa* and *aabb* are not). Prove that if a word has the property that swapping any two adjacent letters makes the word repetitive, then all its letters are identical. (Note that one may swap two adjacent identical letters, leaving a word unchanged.)

Video

https://youtu.be/3Bgka1mOzq0

External Link

https://aops.com/community/p2659401

Solution

A word with the property that swapping any two adjacent letters makes the word repetitive, but for which all letters the same, will called a *unicorn*. The problem asks to show there are no unicorns.

So assume for contradiction there exists a unicorn. We proceed in three steps.

Step 1. Reduction to binary alphabet. In the alphabet, choose one letter (that appears in the unicorn) to be 0 and replace all the letters with 1. The resulting word is still a unicorn on the two symbols $\{0, 1\}$. In this way we only need to consider the problem for the alphabet $\Sigma = \{0, 1\}$.

Step 2. Unicorns are repetitive themselves. So let $w \in \Sigma^n$ be a unicorn. It's easy to see that $w = 010101\cdots$ is not a unicorn; if we swap the first two fits, the resulting word $100101\cdots$ has a doubled 00 that appears only once, so it is not a unicorn. Similarly, $w = 101010\cdots$ is not a unicorn. Hence, we may assume our unicorn w has at least two adjacent bits equal. This means that w is itself repetitive.

Step 3. Main argument. From now on let $w[i]$ denote the i^{th} letter of w. Fix an index $1 \le k < n/2$ for which $w[k] \neq w[k+1]$. Let w' be the word that results from swapping the unequal bits, so $w'[k] = w[k+1]$ and $w'[k+1] = w[k]$. Now w is repetitive, and so is w'; so there should exist integers $p \le n/2$ and $q \le n/2$ dividing n (the *periods*) such that the identities

$$w[i] = w[i+p] \quad \forall 1 \le i \le n-p; \qquad w'[i] = w'[i+q] \quad \forall 1 \le i \le n-q.$$

We may assume p and q are different and don't divide each other. We will now produce a contradiction in two cases.

- Consider the case $\gcd(p, q) = 1$. Then, the number of 1's in w ought to be a multiple of n/p, because w consists of n/p repetitions of some word. Similarly, the number of 1's in w' ought to be a multiple of n/q. But these numbers are equal. So the number of 1's is a multiple of both n/p and n/q. When $\gcd(p, q) = 1$, this can only occur if the numbers 1's is divisible by n, i.e. $w = 0\dots0$ or $w = 1\dots1$. Then w is not in fact a unicorn.

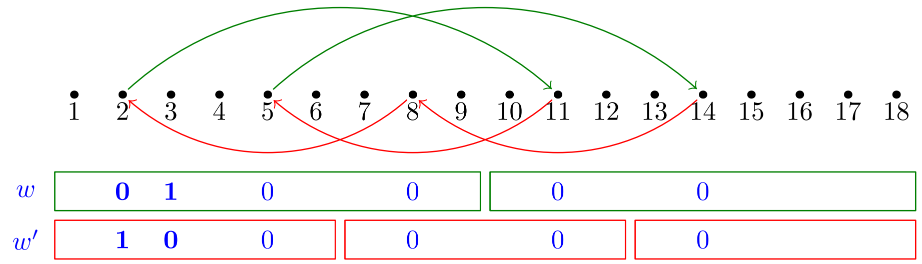
- Otherwise, let $d = \gcd(p, q) > 1$. Choose the index $i = k$ to start. In a *move* we update the index as follows:
  - If $i + p < n$, replace i with $i + p$.
  - Otherwise, replace i with $i - q$.
  - Repeat this until $i \equiv k \pmod q$ again for the first time.

  This process is well defined since p and q are different and less than $n/2$. This creates a sequence

  $$k = i_0, i_1, i_2, \dots, i_m$$

  with the property that the values of w and w' at all indices except the ends are equal. (The assumption $d > 1$ means that all $i_\bullet$ are $k \pmod d$, so this sequence never contains $k + 1$.) Following the chain then creates a contradiction, because we started at $w[k]$ and ended at $w'[i_m] = w'[k] \neq w[k]$ following only equalities.

A cartoon is shown below for $p = 9$, $q = 6$, $n = 18$, $k = 2$.



Remark (The smart solution)**.** An alternative clever way to do this last part is using roots of unity. The idea is to show that in our binary language, if S is the subset of indices with 1's among $\{1, \dots, n\}$, then every repetitive word satisfies

$$\sum_{s \in S} \exp\left(\frac{2\pi n}{s}\right) = 0.$$

(The converse is not true.) Hence a repetitive unicorn cannot exist; swapping *any* two adjacent indices with unequal bits makes the sum change. (The proof earlier shows this too, since we only used one value of k.)