2017 TSTST #2

Tristan Shin

24 Jun 2017

Ana and Banana are playing a game. First Ana picks a word, which is defined to be a nonempty sequence of capital English letters. (The word does not need to be a valid English word.) Then Banana picks a nonnegative integer k and challenges Ana to supply a word with exactly k subsequences which are equal to Ana's word. Ana wins if she is able to supply such a word, otherwise she loses.

For example, if Ana picks the word "TST", and Banana chooses $k = 4$, then Ana can supply the word "TSTST" which has 4 subsequences which are equal to Ana's word.

Which words can Ana pick so that she wins no matter what value of k Banana chooses?

(The subsequences of a string of length n are the 2^n strings which are formed by deleting some of its characters, possibly all or none, while preserving the order of the remaining characters.)

Some notation:

- Addition with words will mean concatenation (so if ω_1 and ω_2 are words, then $\omega_1 + \omega_2$ represents the word with ω_1 followed by ω_2)
- Scalar multiplication by a nonnegative integer will mean repeating the word (so if ω_1 is a word, then $3\omega_1 = \omega_1 + \omega_1 + \omega_1$)
- Magnitude of a word denotes length (so if ω_1 is a word, then $|\omega_1|$ denotes the number of letters in ω_1)

Let $\mathcal{S}$ be the set of words ω such that for any letter ℓ in ω, either one of the neighbors of ℓ is the same letter as ℓ. For example, AABB is in $\mathcal{S}$, but not TSSAL, MOP, A, or TSTST.

The answer is any word <u>not</u> in $\mathcal{S}$.

Let ω be a word not in $\mathcal{S}$. Then we can write

$$\omega = \omega_1 + \ell + \omega_2, |\ell| = 1$$

where neither neighbor of ℓ is the same letter as ℓ. Let the letter in ℓ be $\mathcal{L}$. Then if Banana chooses $k = n$, then Ana can give

$$\omega_1 \underbrace{\mathcal{L}\mathcal{L}\ldots\mathcal{L}}_{n} \omega_2.$$

In order to preserve ω, no letters in ω_1 and ω_2 can be removed. Thus, we must remove $n-1$ letters from the n $\mathcal{L}$'s. There are $\binom{n}{n-1} = n$ ways to do this. All of them bring back ω, so this works.

Now, let ω be a word in $\mathcal{S}$. We will prove that it is impossible that Ana can win $k = 2$. Assume that Ana can supply a word ω' when Ana chooses $k = 2$.

Let $\mathcal{C}$ be the set of letters in ω. Clearly, ω is a substring of ω', so we can say that we are expanding on ω (and adding letters to ω) in order to form ω'. Picking any letter which is not in $\mathcal{C}$ to add to ω does not help nor hurt, as it must be deleted in any formation of ω. Thus, we may assume that ω' consists only of letters in $\mathcal{C}$.

Let

$$\omega = a_1\ell_1 + a_2\ell_2 + \ldots + a_k\ell_k$$

for some letters $\ell_1, \ell_2, \ldots, \ell_k$ and positive integers $a_1, a_2, \ldots, a_k$. Let the "ℓ_i block" represent the block of a_i ℓ_i's inside ω.

Assume that we add a letter $\mathcal{L} \in \mathcal{C}$ to ω in order to form ω'. Let $\mathcal{L} = \ell_M$ for some index M. I claim that one of the following occurs:

1) $\mathcal{L}$ is deleted in any formation of ω from ω', in which case $\mathcal{L}$ is useless and we can ignore it. 2) $\mathcal{L}$ is a neighbor to the ℓ_M block.

Assume that $\mathcal{L}$ is successful in forming a ω from ω', so it is not deleted. Furthermore assume that $\mathcal{L}$ is not a neighbor to the ℓ_M block. We can assume that $\mathcal{L}$ is to the left of the ℓ_M block, otherwise we can flip over the word. Let the letter immediately to the left of $\mathcal{L}$ be ℓ_K, $K < M$. Then if $K + 1 < M$, there must exist the subsequence

$$\phi = a_{K+1}\ell_{K+1} + a_{K+2}\ell_{K+2} + \ldots + a_{M-1}\ell_{M-1}$$

to the left of $\mathcal{L}$ in ω' in order for $\mathcal{L}$ to be useful. We must also append the remainder of the ℓ_K block to the right of $\mathcal{L}$ before ϕ in order for $\mathcal{L}$ to be useful. Either way, we must[hide=*]We might add on more letters, but this can only increase the number of subsequences of ω' that are equal to ω as we can just delete them to get this word.[/hide] have something in the form

$$\omega_1 + \mathcal{L} + \omega_2 + a_M\ell_M + \omega_3,$$

where

$$\omega_1 + a_M\ell_M + \omega_3 = \omega.$$

Then we must delete all of ω_2 and 1 of the ℓ_M's. But there are

$$a_M + 1 \geq 3$$

of them, so there are at least 3 subsequences of ω' that are equal to ω, contradiction. Thus, either 1 or 2 occurs.

If 1 occurs, then $\mathcal{L}$ does not matter and we might as well have deleted it to begin with without affecting the number of subsequences of ω' that are equal to ω.

If 2 occurs, then there are at least

$$a_M + 1 \geq 3$$

ways to remove a ℓ_M, so there are at least 3 subsequences of ω' that are equal to ω, contradiction.

Thus, we cannot place any letter into ω to build ω'. Thus, $\omega' = \omega$. But there is only one subsequence of ω that is equal to ω, so $k = 2$ still loses, contardiction.

Thus, ω' does not exist and thus if ω is in $\mathcal{S}$, then Ana loses on $k = 2$.

Thus, Ana can pick any word not in $\mathcal{S}$ to win no matter what value of k Banana chooses, and picking any word in $\mathcal{S}$ fails. ■