Analytic Combinatorics Homework 6 Problem 2

Eric Neyman
3/14/2017

5/5

Let p be the string we are looking for, i.e. "The quick brown fox jumped over the lazy dog". Observe that p has no autocorrelations (except the trivial one): there is no final substring of the phrase that matches an initial substring. Thus, the autocorrelation polynomial of the string is just $c_p = 1$.

Let S_p be the class of strings where p does not appear as a substring, and T_p be the class of strings ending in p where p does not appear at any point earlier. The solution on Slide 18 generalizes, except that we are no longer working with binary strings, but instead are working with strings where each character has 32 possibilities. Our first construction is therefore $S_p + T_p = E + S_p \times \{Z_0 + Z_1 + \cdots + Z_{31}\}$ (where the Z_i are the characters in our 32-character alphabet). This changes our OGF to $S_p(z) + T_p(z) = 1 + 32zS_p(z)$, which ultimately changes the solution to

$$S_p(z) = \frac{c_p(z)}{z^P + (1 - 32z)c_p(z)}.$$

In particular, $c_p(z) = 1$, and $P = 44$ (counting spaces). Thus, our generating function is

$$S_p(z) = \frac{1}{z^{44} - 32z + 1}.$$

Now, consider $S_p\left(\frac{1}{32}\right)$. This is the sum for all n of the number of strings of length n not containing p, divided by 32^n. Equivalently, it is the sum over all n of the probability that the first n bits of a random string do not contain p, i.e. the sum over all n of the probability that the position of the end of the first occurrence of p (in a random infinite string) is greater than n. But this is simply the expected position of the end of the first occurrence of p. Thus, our answer is simply

$$S_p\left(\frac{1}{32}\right) = \frac{1}{\left(\frac{1}{32}\right)^{44} - 32 \cdot \frac{1}{32} + 1} = \boxed{32^{44} \approx 1.685 \times 10^{66}}.$$