Counting Palindromes According to r-Runs of Ones Using Generating Functions

Helmut Prodinger
Department of Mathematics
Stellenbosch University
7602 Stellenbosch
South Africa
hproding@sun.ac.za

Abstract

Generating functions are derived for the enumeration of all palindromic binary strings of length n having only runs of 1's of length $\le r$. It is demonstrated how one can get asymptotic expressions for fixed r and $n \to \infty$. Eventually, r is treated as a random variable and an asymptotic equivalent for the largest run of 1's in binary palindromes is derived.

1 Enumeration

In the recent paper [3] the interest was in words over the alphabet $\{0, 1\}$ which are *palindromes* and have runs of 1's of bounded length. We firmly believe that *generating functions* are the most appropriate tool here, and since they were not used in [3], we present this natural approach and show as well how one can deal with the case that the maximal 1-run length is treated as a *random variable*. It is worthwhile to note that all our methods can be found in [1].

Let us start with palindromes of *even length*; they are given as ww^R, with a reversed copy of w attached to w. In unrestricted words, the following factorization is appropriate:

$$(\mathbf{0} + \mathbf{1})^* = (\mathbf{1}^*\mathbf{0})^*\mathbf{1}^*.$$

Here, we used the $*$-operation, common in the study of formal languages, so L^* denotes all words that can be formed from concatenating words taken from L in all possible ways. In [1], the notation $\textsc{Seq}(L)$ is mostly used, describing all SEQuences (aka words), formed from L. Now, the mentioned factorization is a very common one for binary words. Each word is (uniquely) decomposed according to each appearence of the letter 0; between them, there are runs (possibly empty) of the letter 1. If a word has s letters 0, then there are $s+1$ such runs of 1's. In terms of generating functions, since the transition $A \to A^*$ means $f \to \frac{1}{1-f}$, the factorization reads as

$$\frac{1}{1-2z} = \frac{1}{1-\dfrac{z}{1-z}}\frac{1}{1-z}.$$

This factorization can immediately be generalized to the instance when the 1-runs should not exceed the parameter r. Then we first consider the set of restricted runs

$$\mathbf{1}^{\le r} = \{\varepsilon, 1, 11, \dots, 1^r\},$$

which translates into

$$1 + z + \cdots + z^r = \frac{1-z^{r+1}}{1-z}.$$

Then we get the formal expression

$$(\mathbf{1}^{\le r}\mathbf{0})^*\mathbf{1}^{\le r},$$

which translates into

$$\frac{1}{1-\dfrac{z(1-z^{r+1})}{1-z}}\frac{1-z^{r+1}}{1-z} = \frac{1-z^{r+1}}{1-2z+z^{r+2}}.$$

Now, going to palindromes of even length, the last group of ones must be bounded by $\lfloor \frac{r}{2} \rfloor$. So a syntactic description of palindromes of even length with bounded 1-runs is

$$(\mathbf{1}^{\le r}\mathbf{0})^*\mathbf{1}^{\le \lfloor \frac{r}{2} \rfloor};$$

this describes the first half of the word only. From this we go immediately to generating functions, by replacing both letters by a variable z. In this way, we count half of the length of the palindromes of even length. If one wants the full length, one must replace z by z^2. So we get

$$\frac{1}{1-z\dfrac{1-z^{r+1}}{1-z}}\frac{1-z^{\lfloor \frac{r}{2} \rfloor+1}}{1-z} = \frac{1-z^{\lfloor \frac{r}{2} \rfloor+1}}{1-2z+z^{r+2}}. \tag{1}$$

One can read off the coefficient of z^n in the power series expansion of this expression, which leads to a clumsy expression:

Set

$$a_{n,r} = [z^n]\frac{1}{1-2z+z^r},$$

then

$$a_{n,r} - 2a_{n-1,r} + a_{n-r,r} = 0,$$

and initial conditions $a_{n,r} = 2^n$ for $n < r$.

Then the number of palindromes of even length $2n$ with all 1-runs $\le r$ is given by

$$[z^n]\frac{1 - z^{\lfloor \frac{r}{2} \rfloor + 1}}{1 - 2z + z^{r+2}} = a_{n,r+2} - a_{n-\lfloor \frac{r}{2} \rfloor - 1, r+2}.$$

We can alternatively express the coefficients in (1) using the *higher order Fibonacci numbers*, as it was done in [3]: Consider $U_{n,r} = U_{n-1,r} + \cdots + U_{n-r,r}$ for $n \ge r$, with initial values $U_{0,r} = \cdots = U_{r-2,r} = 0$, $U_{r-1,r} = 1$. Then

$$\sum_{n\ge 0} U_{n,r} z^n = \frac{z^{r-1}}{1 - (z^r + \cdots + z)} = \frac{z^{r-1}}{1 - z\frac{1-z^r}{1-z}} = \frac{z^{r-1}(1-z)}{1 - 2z + z^{r+1}}.$$

Further,

$$\sum_{n\ge 0} (U_{0,r} + \cdots + U_{n,r}) z^n = \frac{z^{r-1}}{1 - 2z + z^{r+1}},$$

or

$$\sum_{k=0}^{n+r} U_{k,r+1} = [z^n]\frac{1}{1 - 2z + z^{r+2}}.$$

Consequently

$$\begin{aligned}
[z^n]\frac{1 - z^{\lfloor \frac{r}{2} \rfloor + 1}}{1 - 2z + z^{r+2}} &= [z^n]\frac{1}{1 - 2z + z^{r+2}} - [z^{n-\lfloor \frac{r}{2} \rfloor - 1}]\frac{1}{1 - 2z + z^{r+2}} \\
&= \sum_{k=0}^{n+r} U_{k,r+1} - \sum_{k=0}^{n-\lfloor \frac{r}{2} \rfloor - 1 + r} U_{k,r+1} \\
&= \sum_{k=n-\lfloor \frac{r}{2} \rfloor + r}^{n+r} U_{k,r+1}.
\end{aligned}$$

This is the expression given in [3] once one changes the index of summation. Note that $r - \lfloor \frac{r}{2} \rfloor = \lceil \frac{r}{2} \rceil$.

Now we move to palindromes of odd length with middle letter 1: $w1w^R$. Then w is described by

$$(\mathbf{1}^{\le r}\mathbf{0})^* \mathbf{1}^{\le \lfloor \frac{r-1}{2} \rfloor}.$$

In this way, the last group of ones plus the middle 1 plus the first group of ones of the reversed word is still $\le r$ as it should.

The corresponding generating function is

$$\frac{1 - z^{\lfloor \frac{r-1}{2} \rfloor + 1}}{1 - 2z + z^{r+2}}$$

and the coefficient of z^n (counting palindromes of odd length $2n+1$ with middle letter 1) is

$$a_{n,r+2} - a_{n-\lfloor\frac{r-1}{2}\rfloor-1,r+2}.$$

Again, we can alternatively express the corresponding number by higher order Fibonacci numbers:

$$\begin{aligned}[z^n]\frac{1-z^{\lfloor\frac{r-1}{2}\rfloor+1}}{1-2z+z^{r+2}} &= [z^n]\frac{1}{1-2z+z^{r+2}} - [z^{n-\lfloor\frac{r-1}{2}\rfloor-1}]\frac{1}{1-2z+z^{r+2}}\\ &= \sum_{k=0}^{n+r} U_{k,r+1} - \sum_{k=0}^{n-\lfloor\frac{r-1}{2}\rfloor-1+r} U_{k,r+1}\\ &= \sum_{k=n-\lfloor\frac{r-1}{2}\rfloor+r}^{n+r} U_{k,r+1}.\end{aligned}$$

Note that $r-\lfloor\frac{r-1}{2}\rfloor = \lfloor\frac{r}{2}\rfloor+1$.

Finally we move to palindromes of odd length with middle letter 0: $w0w^R$. Then we have

$$(\mathbf{1}^{\le r}\mathbf{0})^*\mathbf{1}^{\le r},$$

since the middle 0 interrupts the last run of ones of the first group. The corresponding generating function is

$$\frac{1-z^{r+1}}{1-2z+z^{r+2}},$$

where again the coefficient of z^n refers to a palindrome of length $2n+1$ with middle 0.

Explicitly we get

$$[z^n]\frac{1-z^{r+1}}{1-2z+z^{r+2}} = a_{n,r+2} - a_{n-r-1,r+2}.$$

In terms of higher order Fibonacci numbers, this reads

$$\begin{aligned}[z^n]\frac{1-z^{r+1}}{1-2z+z^{r+2}} &= [z^n]\frac{1}{1-2z+z^{r+2}} - [z^{n-r-1}]\frac{1}{1-2z+z^{r+2}}\\ &= \sum_{k=0}^{n+r} U_{k,r+1} - \sum_{k=0}^{n-1} U_{k,r+1}\\ &= \sum_{k=n}^{n+r} U_{k,r+1}.\end{aligned}$$

2 Asymptotics

We refer to the paper [2] which might be the first to consider asymptotics for words of restricted runs. The recent paper [4] has many examples of this type. Here, we only consider the key steps and refer for error bounds to the cited literature.

One has to study the *dominant* zero of the denominator, denoted by ρ, which is close to $\frac{1}{2}$ when r gets large (no restriction). From

$$1-2\rho+\rho^{r+2}=0$$

we infer

$$\rho=\frac{1}{2}+\frac{1}{2}\rho^{r+2}\approx\frac{1}{2}+\frac{1}{2^{r+3}}.$$

This procedure is called *bootstrapping*. We also need the constant A in

$$\frac{1}{1-2z+z^{r+2}}\sim\frac{A}{1-z/\rho}\qquad\text{as } z\to\rho,$$

which we get by L'Hopital's rule as

$$A=\frac{-1/\rho}{-2+(r+2)z^{r+1}}\bigg|_{z=\rho}=\frac{-1/\rho}{-2+(r+2)\rho^{r+1}}=\frac{1}{2\rho-(r+2)(2\rho-1)}.$$

So we get the following asymptotic formulæ, valid for $n\to\infty$ and fixed r:

$$[z^n]\frac{1-z^{\lfloor\frac{r}{2}\rfloor+1}}{1-2z+z^{r+2}}\sim(1-\rho^{\lfloor\frac{r}{2}\rfloor+1})A\rho^{-n},$$
$$[z^n]\frac{1-z^{\lfloor\frac{r-1}{2}\rfloor+1}}{1-2z+z^{r+2}}\sim(1-\rho^{\lfloor\frac{r-1}{2}\rfloor+1})A\rho^{-n},$$
$$[z^n]\frac{1-z^{r+1}}{1-2z+z^{r+2}}\sim(1-\rho^{r+1})A\rho^{-n}.$$

And now we turn to the instance where r is a random variable X, and compute, as a showcase, the expected value, so we answer the question about the average value of the longest 1-run in palindromes, in the 3 respective models. As mentioned, this was basically done already by Knuth. When r gets large, the constant A may be replaced by 1, terms of the form ρ^r may be dropped, and in ρ^{-n}, it is enough to use the approximation

$$\rho^{-n}\sim 2^n\big(1-2^{-r-2}\big)^n\sim 2^n\exp(-n/2^{r+2}).$$

Furthermore, to get a probability distribution, we have to divide by 2^n, which is the number of binary words of length n. So the probability that the parameter X is $\le r$ is in all 3 instances approximated by

$$\exp(-n/2^{r+2}).$$

For an expected value, one has to compute

$$\sum_{r\ge0}\big[1-\exp(-n/2^{r+2})\big].$$

This evaluation can be found in many texts ([2, 1, 4]); it is done with the *Mellin transform*, and the result is

$$\log_2 n + \frac{\gamma}{\log 2} - \frac{3}{2} - \frac{1}{\log 2} \sum_{k \neq 0} \Gamma\Big(\frac{2k\pi i}{\log 2}\Big) e^{-2k\pi i \cdot \log_2 n}.$$

Observe that the series in this expression represents a periodic function with small amplitude.

Asymptotically, thus, palindromes with middle letter 0 resp. 1 resp. no middle letter all lead to the same result.

References

[1] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, 2009.

[2] D. E. Knuth. The average time for carry propagation. *Indagationes Mathematicae*, 40:238–242, 1978.

[3] M. A. Nyblom. Counting palindromic binary strings without r-runs of ones. *Journal of Integer Sequences*, 8:8 pages, 2013.

[4] H. Prodinger and S. Wagner. Bootstrapping and double-exponential limit laws. *submitted*, xx:xxx–xxx, 2012.