Symmetric Subsequences from the Left

Jack Hanke

May 15, 2024

Consider a sequence $s = (s_1, \dots, s_n)$ where $s_i \in \mathbb{Z}_b$. We are interested in the function $f : s \to \mathbb{Z}^+$ defined as follows. Find the largest i so that $(s_1, s_2, \dots, s_i)$ is symmetric. Continue by finding the largest j so that $(s_{i+1}, s_{i+2}, \dots, s_j)$ is symmetric. Repeat until s_n is part of a symmetric subsequence. Also let g return the sequence of lengths of symmetric subsequences in s. For example, $f((1,1,0,1,0,0)) = 3$ and $g((1,1,0,1,0,0)) = (2,3,1)$, as the process yields $(1,1)$, $(0,1,0)$, and (0).

Next, let the addition of sequences $s = (s_1, \dots, s_n)$ and $t = (t_1, \dots, t_m)$ to be $s + t = (s_1, \dots, s_n, t_1, \dots, t_m)$. Note in general that $f(s+t) \neq f(t+s)$. Finally, let

$$m_{b,n} = \max_{s,b} f(s).$$

Theorem 1. *$m_{b,n} = n$ for $b > 2$*

Proof. Construct the sequence $s = (s_1, \dots, s_n)$ with $s_i = i \mod b$. This sequence has the maximum number of symmetric subsequences n. First, notice the values equivalent to s_1 are s_{bk+1} for integer k so that $bk + 1 \leq n$, and these values are the only candidates to check for possible symmetric subsequences for s_1. However, the subsequence between s_{bk+1} and $s_{b(k+1)+1}$ cannot be symmetric for $b > 2$, as $(i+1) \mod b \neq (i-1) \mod b$ unless $b = 2$. As the $(i+1) \mod b \neq (i-1) \mod b$ holds for all i, every s_i must be a member of a length 1 symmetric subsequence. ☐

Now consider the $b = 2$ case.

Lemma 1. *A length 1 symmetric subsequence in position i implies $s_{i+1} = s_{i+2} = \cdots = s_n$.*

Proof. Suppose there exists a length 1 symmetric subsequence in position i. Then suppose there exists a value j greater than i so that $s_j = s_i$. Therefore $(s_{i+1}, \dots, s_{j-1})$ must all be the opposite value to s_i. However, $(s_i, \dots, s_j)$ form a symmetric subsequence. Therefore there can exist no such j. ☐

Lemma 2. *There cannot exist three consecutive pairs of length 2 symmetric sequences.*

Proof. Suppose there exists a subsequence $s_i, s_{i+1}, s_{i+2}, s_{i+3}, s_{i+4}, s_{i+5}$ where the three consecutive pairs form symmetric subsequences. This implies that $s_i = s_{i+1}$, $s_{i+2} = s_{i+3}$, and $s_{i+4} = s_{i+5}$. As (s_i, s_{i+1}) form a length 2 symmetric subsequence, $s_i \neq s_{i+2}$, as otherwise $(s_i, s_{i+1}s_{i+2}, s_{i+3})$ would form a larger symmetric subsequence. Similarly, $s_{i+2} \neq s_{i+4}$. However, this gives $s_i = s_{i+4}$, which implies $(s_i, s_{i+1}s_{i+2}, s_{i+3}, s_{i+4}, s_{i+5})$ forms a symmetric sequence. ☐

Theorem 2. $m_{2,n} \leq \lfloor \frac{3n+10}{7} \rfloor$

Proof. The first seven values for $m_{2,n}$ are $1, 2, 2, 3, 3, 4, 4$. We will build maximal sequences using these base sequences to form the upper bound.

From Lemma **??** we know that the first symmetric subsequences of length 1 always proceeds the final symmetric subsequence. Therefore, the placement of the first length 1 symmetric subsequence to maximize the density of subsequences is the $n-1$ position, which leaves the final sequence with 2 length 1 symmetric subsequences. Therefore a sequence s' that maximizes f could have

$$g(s') = (\ldots, 1, 1).$$

From Lemma **??** we know that there cannot exist 3 consecutive length 2 symmetric subsequences. Therefore s' must also have

$$g(s') = (\ldots, 3, 2, 2, 1, 1).$$

We can continue adding a subsequence t that has $g(t) = (3, 2, 2)$ to the beginning of s' for as long as we'd like. Truncating the start of s' at the beginnings of each symmetric subsequence gives an upper bound for $m_{2,n}$ for $n = 2, 4, 6 \mod 7$. Specifically, for $n = 2 \mod 7$ we have $m_{2,n} \leq m_{2,2} + \frac{3n}{7}$, $n = 4 \mod 7$ we have $m_{2,n} \leq m_{2,4} + \frac{3n}{7}$, and $n = 6 \mod 7$ we have $m_{2,n} \leq m_{2,6} + \frac{3n}{7}$.

To achieve an upperbound for $n = 1, 3, 5 \mod 7$, consider s'', which has

$$g(s'') = (\ldots, 3, 2, 2, 1).$$

Similarly we can continue adding a subsequence t that has $g(t) = (3, 2, 2)$ to the beginning of s'', and truncating the start of s'' at the beginnings of each symmetric subsequence gives the same form of an upperbound for $n = 1, 3, 5 \mod 7$. Finally, for $n = 0 \mod 7$, consider s''', which has

$$g(s''') = (\ldots, 2, 3, 2, 1, 1)$$

For s''' we can continue adding a subsequence t that has $g(t) = (2, 3, 2)$ to the beginning of s''', and truncating the start of s''' at the beginnings of each symmetric subsequence gives the same form of an upperbound for $n = 0, 2, 4 \mod 7$.

This gives us the upper bound $\forall n$.

□

Lemma 3. *The sequence* $s^* = (0, 0, 1, 1, 0, 1, 0)$ *has* $g(s^*) = (2, 2, 3)$ *and*

$$f\left(\sum_{i=1}^{k} s^*\right) = \sum_{i=1}^{k} f(s^*) = 3k$$

$\forall k \geq 1$.

Proof. We must first verify that the largest symmetric subsequence containing s_1 in $S = \sum_{i=1}^{k} s^*$ is (s_1, s_2). Notice that $s_i = 0$ for $i = 4, 6 \mod 7$ cannot be the end of a symmetric subsequence, as $s_i = 1$ for $i = 3, 5 \mod 7$. Similarly $s_i = 2$ for $i = 2 \mod 7$ has $(s_1, s_2, s_3) = (0, 0, 1)$ and $(s_{i-2}, s_{i-1}, s_i) = (0, 0, 0)$ for $i > 2$. Finally, $s_i = 0$ for $i = 1 \mod 7$ has $(s_1, s_2, s_3, s_3) = (0, 0, 1, 1)$ and $(s_{i-3}, s_{i-2}, s_{i-1}, s_i) = (0, 1, 0, 0)$ for $i > 1$. So (s_1, s_2) is the largest symmetric subsequence containing s_1

As s_2 is part of the first subsequence, we next prove the largest symmetric subsequence containing s_3 is (s_3, s_4). For $i = 3 \mod 7$, we have $(s_3, s_4, s_5, s_6) = (1, 1, 0, 1)$ and $(s_{i-3}, s_{i-2}, s_{i-1}, s_i) = (0, 0, 1, 1)$. For $i = 4, 6 \mod 7$, we have $(s_3, s_4) = (1, 1)$ and $(s_{i-1}, s_i) = (0, 1)$.

As s_4 is part of the first subsequence, we next prove the largest symmetric subsequence containing s_5 is (s_5, s_6, s_7). For $i = 1, 2 \mod 7$, we have $(s_5, s_6) = (0, 1)$ and $(s_{i-1}, s_i) = (0, 0)$. For $i = 5 \mod 7$, we have $(s_5, s_6, s_7) = (0, 1, 0)$ and $(s_{i-2}, s_{i-1}, s_i) = (1, 1, 0)$. For $i = 0 \mod 7$, we have $(s_5, s_6, s_7, s_8) = (0, 1, 0, 0)$ and $(s_{i-3}, s_{i-2}, s_{i-1}, s_i) = (1, 0, 1, 0)$.

□

Lemma 4. *The sequence $s^{**} = (0, 0, 1, 0, 1, 1, 1)$ has $g(s^*{*}) = (2, 3, 2)$ and*

$$f\left(\sum_{i=1}^{k} s^{**}\right) = \sum_{i=1}^{k} f(s^{**}) = 3k$$

$\forall k \geq 1$.

Proof. Proof mirrors the structure of Lemma **??**. □

Theorem 3. $m_{2,n} \geq \lfloor \frac{3n+10}{7} \rfloor$

Proof. To prove the lower bound we simply construct sequences that satisfy the bound for certain residues mod 7.

First, let $s^* = (0, 0, 1, 1, 0, 1, 0)$. Then let k be the smallest multiple of 7 greater than or equal to n, and construct the sequence

$$S = \sum_{i=1}^{k} s^*.$$

Truncating S after $S_{n-6}, S_{n-4}, S_{n-2}$, and S_{n-1} gives a sequence of length congruent to $i = 1, 3, 5, 6 \mod 7$ and contains $\frac{3}{7}(k - 7) + m_{2,i}$ symmetric subsequences.

For $n = 0 \mod 7$ construct

$$T = (0) + \sum_{i=1}^{k} s^*.$$

We must verify that the largest symmetric subsequence containing t_1 is (t_1, t_2, t_3). For $i = 1, 6 \mod 7$ $(t_1, t_2) = (0, 0)$ and $(t_i, t_{i+1} = (1, 0))$. For $i = 2 \mod 7$ $(t_1, t_2, t_3) = (0, 0, 0)$ and $(t_i, t_{i+1}, t_{i+2} = (1, 0, 0))$. For $i = 3 \mod 7$ $(t_1, t_2, t_3, t_4, t_5) = (0, 0, 0, 1, 1)$ and $(t_i, t_{i+1}, t_{i+2}, t_{i+3}, t_{i+4}) = (0, 1, 0, 0, 0)$. Therefore, (t_1, t_2, t_3) is the largest symmetric subsequence containing t_1 in T. For the rest of the sequence, as the subsequence (t_4, t_n) is the same as s_3, s_n, the remainder of the proof of Lemma **??** holds.

Truncating T at T_{n-1} gives a sequence of length congruent to $0 \mod 7$ and contains $\frac{3}{7}(k-7)+m_{2,7}$ symmetric subsequences.

For $n=2 \mod 7$, let k be the smallest multiple of 7 less than or equal to n, and construct the sequence

$$U=\sum_{i=1}^{k} s^{*}+(0,1).$$

The addition of 0 has no affect on the symmetric runs from Lemma **??**, but the addition of 1 must be examined. As only the positions $i=2 \mod 7$ is the beginning of a symmetric subsequence in U that starts with 1, we only need to consider U_i where $i=2 \mod 7$. We have $(U_i, U_{i+1})=(1,1)$ and $(U_{n-1}, U_n)=(0,1)$. The sequence U gives a sequence of length congruent to $2 \mod 7$ and contains $\frac{3}{7}(k-7)+m_{2,2}$ symmetric subsequences.

For $n=4 \mod 7$, we use $s^{**}=(0,0,1,0,1,1,1)$, and let k be the smallest multiple of 7 greater than or equal to n, and construct the sequence

$$V=\sum_{i=1}^{k} s^{**}$$

Truncating V after V_{n-3} gives a sequence of length congruent to $4 \mod 7$ and contains $\frac{3}{7}(k-7)+m_{2,4}$ symmetric subsequences.

$\square$

Therefore, $m_{2,n}=\lfloor\frac{3n+10}{7}\rfloor$.