AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **84(3)** (2022), Pages 419–430

Intertwining of complementary Thue-Morse factors

Jeffrey Shallit

School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1
Canada
shallit@uwaterloo.ca

Abstract

Generalizing a 2009 question of Bernhardt, we consider the positions of occurrences of a factor x and its binary complement $\overline{x}$ in the Thue-Morse word $\mathbf{t} = \mathtt{01101001}\cdots$, and show that these occurrences are "intertwined" in essentially just two different ways. Our proof method consists of stating the needed properties as a first-order logic formula φ, and then using a theorem-prover to prove φ.

1 Introduction

The Thue-Morse sequence $\mathbf{t} = t_0 t_1 t_2 \cdots = \mathtt{01101001}\cdots$ is a famous binary sequence with many interesting properties [1]. In this note we prove yet another in a long list of such properties, this time concerning complementary factors.

A *factor* of an infinite word $\mathbf{w}$ is a contiguous block sitting inside $\mathbf{w}$. In this paper we will only be concerned with factors of finite length. Define $\overline{\mathtt{0}} = \mathtt{1}$ and $\overline{\mathtt{1}} = \mathtt{0}$, and extend this notion to words in the obvious way, so that if $w = a_1 a_2 \cdots a_n$, then $\overline{w} = \overline{a_1}\,\overline{a_2} \cdots \overline{a_n}$. We say two binary words x, y are *complementary* if $x = \overline{y}$. Thus, for example, 0110 and 1001 are complementary. We distinguish between factors and *occurrences of factors*; the latter term refers to a particular starting position within $\mathbf{t}$ and length of the factor. We say two occurrences $\mathbf{t}[i..j]$ and $\mathbf{t}[k..\ell]$ of (possibly different) factors *overlap* if they share a position in common, that is, if either $i \leq k \leq j$ or $k \leq i \leq \ell$.

It is well-known that the Thue-Morse word $\mathbf{t}$ is *recurrent*, that is, every factor that occurs has infinitely many occurrences (first observed by Morse [5]). Further, $\mathbf{t}$ is complement-invariant: if a factor x occurs in $\mathbf{t}$, then so does its binary complement $\overline{x}$. This suggests looking at the positions of the consecutive occurrences of x and $\overline{x}$ in $\mathbf{t}$. Two occurrences of a factor and its complement may, of course, overlap each other, as is the case with 010 and 101. However, two occurrences of the *same* factor in Thue-Morse can never overlap, because the Thue-Morse sequence is overlap-free [11].

ISSN: 2202-3518

For example, let us consider the complementary factors 11 and 00, marking the occurrences of 11 in blue and 00 in orange:

$$\texttt{0110100110010110100 1} \cdots$$

Seeing this, it is natural to conjecture that occurrences of 11 and 00 strictly alternate in $\mathbf{t}$, a conjecture that is not hard to prove and appeared in a 2009 paper of Bernhardt [2].

However, strict alternation, as in this example, is not the only possibility for other factors. If we consider the complementary factors 00110 and 11001 instead, then the occurrences behave differently:

$$\texttt{01101001100101101001011001101001} \cdots .$$

(We use green to indicate where the occurrences overlap.) If we write A for an occurrence of 00110 and B for 11001 an occurrence of its complementary factor, then experiments quickly lead to the conjecture that these factors occur in the repeating pattern $(\texttt{ABBA})^\omega = \texttt{ABBAABBAABBA} \cdots$.

In this paper we prove that the two patterns $(\texttt{AB})^\omega$ and $(\texttt{ABBA})^\omega$ are essentially the only nontrivial possibilities for intertwining of complementary factors.

For a deep study of the gaps between successive occurrences of factors in $\mathbf{t}$, see the recent paper of Spiegelhofer [10].

2 The main theorem

Let x be a finite, nonempty factor of the Thue-Morse word $\mathbf{t}$. Consider all occurrences of x and $\overline{x}$ in $\mathbf{t}$ and list their starting positions in increasing order, writing A for an occurrence of x and B for an occurrence of $\overline{x}$. (An occurrence of x may overlap that of $\overline{x}$.) Call the resulting infinite sequence of A's and B's the *intertwining sequence* of x, and write it as $I(x)$.

The following is our main result.

Theorem 2.1. *The only possibilities for $I(x)$ are as follows:*

1. ABBABAABBAABABBA $\cdots$, *which is the Thue-Morse word itself under the coding* $0 \to \texttt{A}$, $1 \to \texttt{B}$;
2. BAABABBAABBABAAB $\cdots$, *which is the Thue-Morse word itself under the coding* $0 \to \texttt{B}$, $1 \to \texttt{A}$;
3. $(\texttt{AB})^\omega$;
4. $(\texttt{BA})^\omega$;
5. $(\texttt{ABBA})^\omega$;

6. $(\mathtt{BAAB})^\omega$.

Furthermore, possibility 1 only occurs if $x = \mathtt{0}$ *and possibility 2 only occurs if* $x = \mathtt{1}$.

Proof. It is trivial to see the claim for $x = \mathtt{0}$ and $x = \mathtt{1}$. So in what follows, we assume $|x| \geq 2$.

The idea of our proof is to write first-order logic formulas for assertions that imply our desired results, and then use the theorem-prover `Walnut` to prove the results. This is a strategy that has been used many times now (see, e.g., [7]). For more about `Walnut`, see [6, 7].

We describe the formulas in detail for the cases $(\mathtt{AB})^\omega$ and $(\mathtt{ABBA})^\omega$, leaving the other cases to the reader.

To assert that the pattern $(\mathtt{AB})^\omega$ describes the occurrences of x and $\overline{x}$ in $\mathbf{t}$, we create first-order logic formulas asserting each of the following:

(a) one of the two words x and $\overline{x}$ occurs at positions j, k for $j < k$, and furthermore that neither of the two words occurs at any position between j and k. This ensures that j and k mark the starting position of two *consecutive* factors chosen from $\{x, \overline{x}\}$.

(b) if j, k are two positions as in (a), then one must be the position of x, while the other is the position of $\overline{x}$. This forces the consecutive occurrences of the factors to alternate, and hence form either the pattern $(\mathtt{AB})^\omega$ or $(\mathtt{BA})^\omega$.

(c) the first occurrence of either x or $\overline{x}$ in $\mathbf{t}$ is actually an occurrence of x. This, together with (b), forces the pattern to be of the form $(\mathtt{AB})^\omega$.

We specify the word x by giving one of its occurrences, that is, two integers i, n such that $x = \mathbf{t}[i..i+n-1]$.

Here is the meaning of each logical formula we now define.

- $\mathrm{feq}(i, j, n)$ asserts that $\mathbf{t}[i..i+n-1] = \mathbf{t}[j..j+n-1]$;
- $\mathrm{feqc}(i, j, n)$ asserts that $\mathbf{t}[i..i+n-1] = \overline{\mathbf{t}[j..j+n-1]}$;
- $\mathrm{either}(i, j, n)$ asserts that either $\mathbf{t}[i..i+n-1] = \mathbf{t}[j..j+n-1]$ or $\mathbf{t}[i..i+n-1] = \overline{\mathbf{t}[j..j+n-1]}$;
- $\mathrm{consec}(i, j, k, n)$ asserts that $j < k$ and $\mathbf{t}[j..j+n-1] \in \{x, \overline{x}\}$ and $\mathbf{t}[k..k+n-1] \in \{x, \overline{x}\}$, where $x = \mathbf{t}[i..i+n-1]$, but no factor starting in between these two equals either x or $\overline{x}$.
- $\mathrm{ab}(i, j, k, n)$ asserts $\mathbf{t}[j..j+n-1] = x$ and $\mathbf{t}[k..k+n-1] = \overline{x}$, for $x = \mathbf{t}[i..i+n-1]$.
- $\mathrm{first}(i, j, n)$ asserts that $\mathbf{t}[j..j+n-1]$ is the first occurrence of the factor $\mathbf{t}[i..i+n-1]$ in $\mathbf{t}$;

- afirst(i, n) asserts that the first occurrence of the factor $x = \mathbf{t}[i..i+n-1]$ precedes the first occurrence of $\overline{x}$ in $\mathbf{t}$;
- abpat(i, n) asserts that the intertwining sequence of $x = \mathbf{t}[i..i+n-1]$ and $\overline{x}$ is $(\mathtt{AB})^\omega$.
- bapat(i, n) asserts that the intertwining sequence of $x = \mathbf{t}[i..i+n-1]$ and $\overline{x}$ is $(\mathtt{BA})^\omega$.

$$\begin{aligned}
\text{feq}(i,j,n) &:= \forall k\ (k < n) \implies \mathbf{t}[i+k] = \mathbf{t}[j+k] \\
\text{feqc}(i,j,n) &:= \forall k\ (k < n) \implies \mathbf{t}[i+k] \neq \mathbf{t}[j+k] \\
\text{either}(i,j,n) &:= \text{feq}(i,j,n) \vee \text{feqc}(i,j,n) \\
\text{consec}(i,j,k,n) &:= (j < k) \wedge \text{either}(i,j,n) \wedge \text{either}(i,k,n) \wedge \forall l\ (j < l \wedge l < k) \\
&\quad \implies \neg\, \text{either}(i,l,n) \\
\text{ab}(i,j,k,n) &:= \text{feq}(i,j,n) \wedge \text{feqc}(i,k,n) \\
\text{first}(i,j,n) &:= \text{feq}(i,j,n) \wedge \forall k\ (k < j) \implies \neg\, \text{feq}(i,k,n) \\
\text{afirst}(i,n) &:= \forall j,k\ (\text{first}(i,j,n) \wedge \text{feqc}(i,k,n)) \implies j < k \\
\text{abpat}(i,n) &:= (n > 0) \wedge \text{afirst}(i,n) \wedge \forall j,k\ \text{consec}(i,j,k,n) \implies \\
&\quad (\text{ab}(i,j,k,n) \vee \text{ab}(i,k,j,n)) \\
\text{bapat}(i,n) &:= (n > 0) \wedge (\neg\, \text{afirst}(i,n)) \wedge \forall j,k\ \text{consec}(i,j,k,n) \implies \\
&\quad (\text{ab}(i,j,k,n) \vee \text{ab}(i,k,j,n))
\end{aligned}$$

The translation into `Walnut` is

```
def feq "Ak (k<n) => T[i+k]=T[j+k]":
def feqc "Ak (k<n) => T[i+k]!=T[j+k]":
def either "$feq(i,j,n)|$feqc(i,j,n)":
def consec "j<k & $either(i,j,n) & $either(i,k,n) & Al (j<l & l<k)
   => ~$either(i,l,n)":
def ab "$feq(i,j,n) & $feqc(i,k,n)":
def first "$feq(i,j,n) & Ak (k<j) => ~$feq(i,k,n)":
def afirst "Aj,k ($first(i,j,n) & $feqc(i,k,n)) => j<k":
def abpat "(n>0) & $afirst(i,n) & Aj,k $consec(i,j,k,n) =>
   ($ab(i,j,k,n)|$ab(i,k,j,n))":
def bapat "(n>0) & (~$afirst(i,n)) & Aj,k $consec(i,j,k,n) =>
   ($ab(i,j,k,n)|$ab(i,k,j,n))":
```

We now do the same thing for the patterns $(\mathtt{ABBA})^\omega$ and $(\mathtt{BAAB})^\omega$. The one complication is that to assert that the intertwining sequence is $(\mathtt{ABBA})^\omega$, for example, then one must assert that

(a) the first two occurrences of either x or $\overline{x}$ form the pattern `AB`;

(b) three consecutive occurrences of either x or $\overline{x}$ in $\mathbf{t}$ must form the pattern ABB or BBA or BAA or AAB.

Let us prove, by induction on k, that if (a) and (b) both hold, then the first $4k+2$ elements of the intertwining sequence must be $\mathtt{AB}(\mathtt{BAAB})^k$. The base case is $k=0$, and from (a) we know the first two occurrences have code AB. Otherwise assume the result is true for k and we prove it for $k+1$. By induction we know the last two occurrences are coded by AB. By (b) the next four occurrences must be, successively, B, then A, then A, then B. This completes the proof.

We now give the Walnut commands for checking the criteria (a) and (b):

```
def firstc "$feqc(i,j,n) & Ak (k<j) => ~$feqc(i,k,n)":
# j is the first occurrence of the complement of t[i..i+n-1]
def abfirst "Aj,k ($first(i,j,n) & $firstc(i,k,n)) =>
   (j<k & Al (j<l & l<k) => ~$either(i,l,n))":
# first two occurrences of t[i..i+n-1] or complement are of the form AB
def abb "$feq(i,j,n) & $feqc(i,k,n) & $feqc(i,l,n)":
def bba "$feqc(i,j,n) & $feqc(i,k,n) & $feq(i,l,n)":
def baa "$feqc(i,j,n) & $feq(i,k,n) & $feq(i,l,n)":
def aab "$feq(i,j,n) & $feq(i,k,n) & $feqc(i,l,n)":
def abbapat "(n>0) & $abfirst(i,n) & Aj,k,l ($consec(i,j,k,n) &
   $consec(j,k,l,n)) => ($abb(i,j,k,l,n) | $bba(i,j,k,l,n) |
   $baa(i,j,k,l,n) | $aab(i,j,k,l,n))":
def baabpat "(n>0) & (~$abfirst(i,n)) & Aj,k,l ($consec(i,j,k,n) &
   $consec(j,k,l,n)) => ($baa(i,j,k,l,n) | $aab(i,j,k,l,n) |
   $abb(i,j,k,l,n) | $bba(i,j,k,l,n))":
```

Now we are ready to finish the proof of the theorem. First we check that

$$\begin{aligned} I(\mathtt{11}) &= I(\mathbf{t}[1..2]) = (\mathtt{AB})^\omega \\ I(\mathtt{00}) &= I(\mathbf{t}[5..6]) = (\mathtt{BA})^\omega \\ I(\mathtt{101}) &= I(\mathbf{t}[2..4]) = (\mathtt{ABBA})^\omega \\ I(\mathtt{010}) &= I(\mathbf{t}[3..5]) = (\mathtt{BAAB})^\omega, \end{aligned}$$

as follows:

```
eval alloccur "$abpat(1,2) & $bapat(5,2) & $abbapat(2,3) & $baabpat(3,3)":
```

and Walnut returns TRUE.

Next, we check that for all i and all $n \geq 2$, the intertwining sequence of $\mathbf{t}[i..i+n-1]$ is either $(\mathtt{AB})^\omega$, $(\mathtt{BA})^\omega$, $(\mathtt{ABBA})^\omega$, or $(\mathtt{BAAB})^\omega$.

```
eval checkeach "Ai,n (n>=2) => ($abpat(i,n)|$bapat(i,n)|$abbapat(i,n)|
   $baabpat(i,n))":
```

and Walnut returns TRUE.

This completes the proof. □

For factors of length $n = 2$, the only intertwining patterns that occur are $(\mathtt{AB})^\omega$ and $(\mathtt{BA})^\omega$. However, for each $n \geq 3$, we can prove that each of the four patterns actually occurs.

Theorem 2.2. *For every $n \geq 3$, and each of the four patterns $p \in \{\mathtt{AB}, \mathtt{BA}, \mathtt{ABBA}, \mathtt{BAAB}\}$, there is a length-$n$ factor x of $\mathbf{t}$ whose occurrence pattern is p^ω.*

Proof. We use Walnut with the command

```
eval checklen "An (n>=3) => Ei,j,k,l $abpat(i,n) & $bapat(j,n) &
   $abbapat(k,n) & $baabpat(l,n)":
```

and Walnut returns TRUE. □

We can give a relatively simple criterion for when the intertwining sequence is of the form $(\mathtt{ABBA})^\omega$ or $(\mathtt{BAAB})^\omega$:

Theorem 2.3. *The intertwining sequence for a factor x of the Thue-Morse sequence is $(\mathtt{ABBA})^\omega$ or $(\mathtt{BAAB})^\omega$ if and only if every occurrence of x in $\mathbf{t}$ always overlaps some occurrence of $\overline{x}$, either to the left or right.*

Proof. We use Walnut, and just give the Walnut code.

```
def absminus "(i>=j & i=j+k) | (j>=i & j=i+k)":
# true if k = |i-j|
def abbaorbaab "$abbapat(i,n)|$baabpat(i,n)":
def all_overlap_complement "Aj $feq(i,j,n) =>
   Ek,m $absminus(j,k,m) & m<n & $feqc(j,k,n)":
# for all occurrences of T[i..i+n-1], it overlaps
# its complement either to the left or right
eval thm3 "Ai,n (n>=1) => ($abbaorbaab(i,n) <=>
   $all_overlap_complement(i,n))":
```

and Walnut returns TRUE for the last command. □

Finally, we prove a result about the overlapping of a factor x and its complement $\overline{x}$.

Theorem 2.4. *Every length-n factor of $\mathbf{t}$ has some occurrence in $\mathbf{t}$ that overlaps its complement if and only if $n = 2^k + 1$ for $k \geq 1$.*

Proof. We use Walnut again.

```
def some_complement_overlap "Ej,k,m $feq(i,j,n) & $absminus(j,k,m)
   & m<n & $feqc(j,k,n)":
# some occurrence of T[i..i+n-1] overlaps its complement
def testover "Ai $some_complement_overlap(i,n)":
```

The result of the last command is an automaton recognizing the base-2 representations of those n for which every length-n factor of $\mathbf{t}$ has some occurrence in $\mathbf{t}$ that overlaps its complement. This automaton is depicted in Figure 1, and accepts exactly the set 10^*1, proving the result.

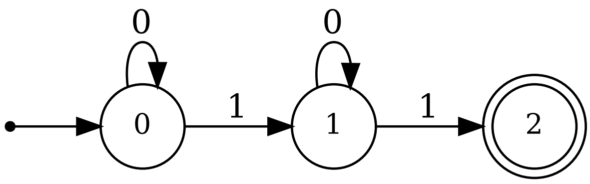

Figure 1: Automaton for Theorem 2.4.

□

3 Automata

We now give the four automata for the four cases of intertwining sequence. Each automaton takes, as input, the base-2 expansions of i and n in parallel, starting with the most significant bit, and accepts if and only if $\mathbf{t}[i..i+n-1]$ has the specified intertwining sequence. The reader can now easily check, for example, in Figure 2, that $[0,1][1,0]$ is accepted, demonstrating that $I(\mathbf{t}[1..2]) = I(\mathtt{11}) = (\mathtt{AB})^\omega$, as we saw in Section 1.

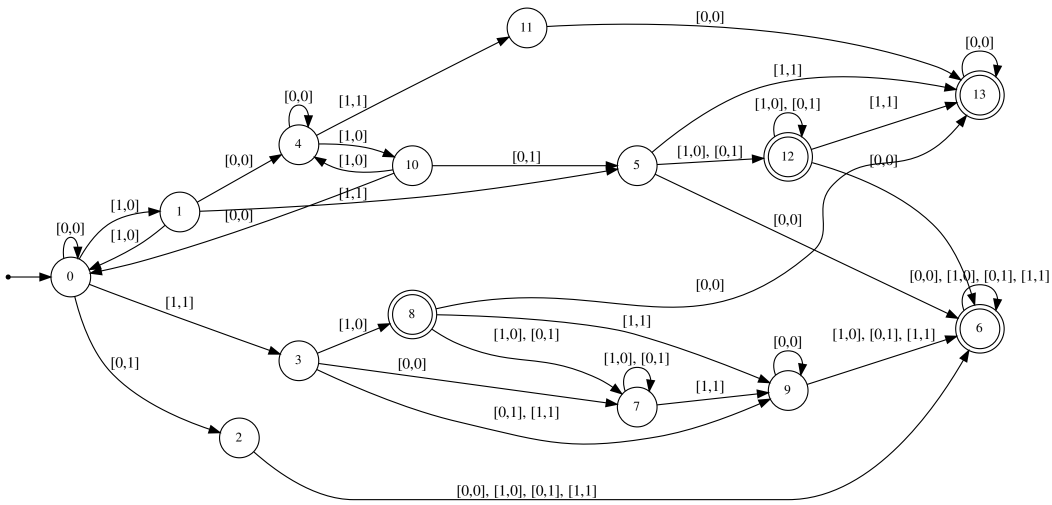

Figure 2: Automaton for (i, n) such that $I(\mathbf{t}[i..i+n-1]) = (\mathtt{AB})^\omega$.

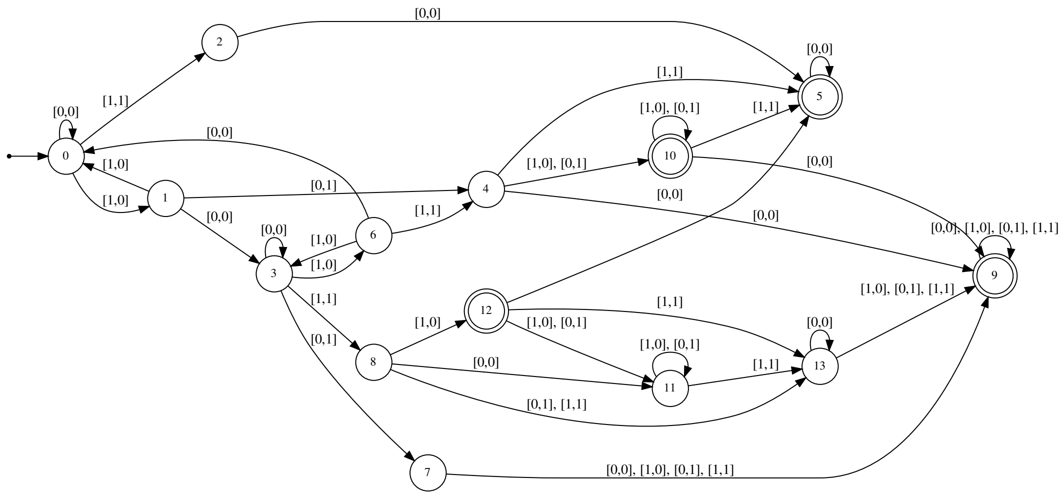

Figure 3: Automaton for (i, n) such that $I(\mathbf{t}[i..i+n-1]) = (\texttt{BA})^\omega$.

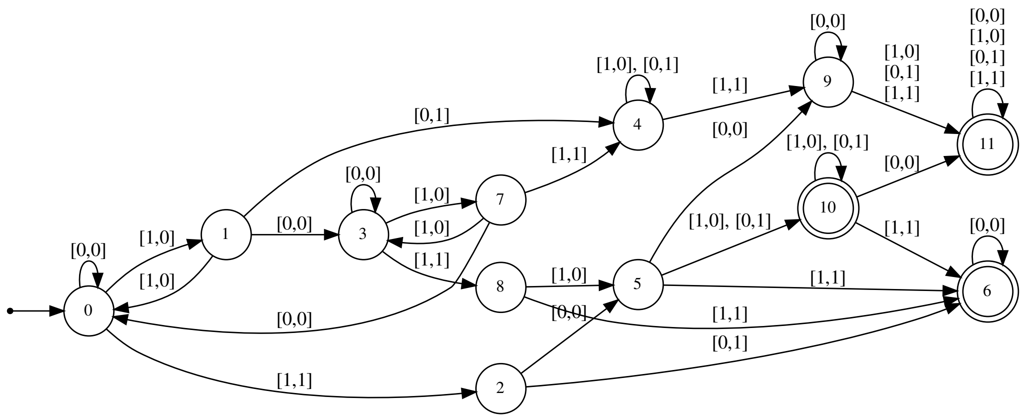

Figure 4: Automaton for (i, n) such that $I(\mathbf{t}[i..i+n-1]) = (\texttt{ABBA})^\omega$.

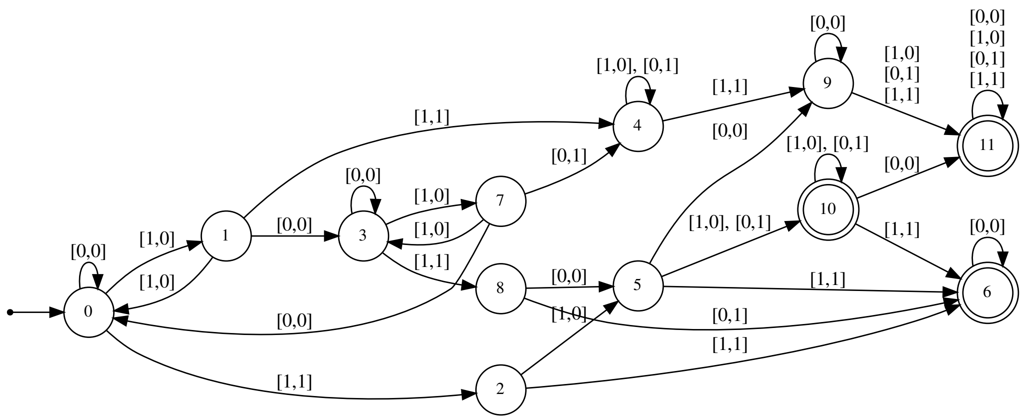

Figure 5: Automaton for (i, n) such that $I(\mathbf{t}[i..i+n-1]) = (\mathtt{BAAB})^\omega$.

Remark 3.1. We can combine these automata, as in [8], to get a single DFAO that, on input (i, n), computes which of the six possibilities in Theorem 2.1 occurs. However, the resulting automaton has 30 states and is rather complicated in appearance, so we do not give it here.

4 Number of factors of each type

We now determine the number of length-n factors of each of the four types. It is easy to see that there is a 1–1 correspondence between length-n factors where the intertwining sequence is $(\mathtt{AB})^\omega$ and those where the intertwining sequence is $(\mathtt{BA})^\omega$, and similarly for those with intertwining sequence $(\mathtt{ABBA})^\omega$ and $(\mathtt{BAAB})^\omega$. Thus it suffices to just handle $(\mathtt{AB})^\omega$ and $(\mathtt{ABBA})^\omega$.

Let $f(n)$ be the number of length-n factors x of $\mathbf{t}$ where $I(x) = (\mathtt{AB})^\omega$, and let $g(n)$ be the number of length-n factors x of $\mathbf{t}$ where $I(x) = (\mathtt{ABBA})^\omega$. Here is a table of the first few values of these functions:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$f(n)$	0	2	2	4	4	6	8	8	8	10	12	14	16	16	16
$g(n)$	0	0	1	1	2	2	2	3	4	4	4	4	4	5	6

Sequence $f(n)$ is A352227 in the OEIS [9] and $g(n)$ is A352228. It turns out that both of these sequences are expressible in terms of two known sequences in the OEIS.

The sequence A006165 is defined as follows:

$$\begin{aligned} \text{A006165}(0) &= \text{A006165}(1) = 1 \\ \text{A006165}(2n) &= 2 \cdot \text{A006165}(n), \quad n \geq 2 \\ \text{A006165}(2n+1) &= \text{A006165}(n+1) + \text{A006165}(n), \quad n \geq 1. \end{aligned}$$

It arises in the so-called "Josephus problem" where n people, numbered $1, 2, \ldots, n$ are arranged in a circle and every second person is marked until only one remains; this person is then removed and the process continues again from the start, until only one remains. The last person removed is $a(n+1)$.

The sequence A060973 is defined as follows:

$$\begin{aligned} \text{A060973}(0) &= \text{A060973}(1) = 0 \\ \text{A060973}(1) &= 1 \\ \text{A060973}(2n) &= 2 \cdot \text{A060973}(n), \quad n \neq 1 \\ \text{A060973}(2n+1) &= \text{A060973}(n+1) + \text{A060973}(n), \quad n \geq 0. \end{aligned}$$

Both of these sequences are examples of "divide-and-conquer" recurrences that frequently arise in the analysis of algorithms; see [4].

Theorem 4.1. *We have*

$$\begin{aligned} f(n+1) &= 2 \cdot \text{A006165}(n) \quad \textit{for } n \geq 1; \\ g(n+1) &= \text{A060973}(n) \quad \textit{for } n \geq 0, \end{aligned}$$

where the sequence numbers refer to sequences in the On-Line Encyclopedia of Integer Sequences (OEIS) [9].

Proof. Let us start with $(\texttt{AB})^\omega$. Using the `Walnut` commands

```
def firstocc "Aj (j<i) => ~$feq(i,j,n)":
eval mab n "$firstocc(i,n+1) & $abpat(i,n+1)":
```

we can construct the so-called "linear representation" for $f(n+1)$. Such a representation expresses a function in the form $v\gamma(x)w$, where x is the binary representation of n, starting with the most significant digit. Here v is a row vector of size k, w is a column vector of size k, and $\gamma(x)$ is a $k \times k$-matrix-valued morphism obeying the product rule $\gamma(yz) = \gamma(y)\gamma(z)$ for all strings y, z. For more about these representations, see [3]. The rank of a linear representation is the number k; a representation is minimal if no linear representation of smaller rank represents the same function.

The linear representation for $f(n+1)$ is computed as

$$v_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \qquad \gamma_1(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \gamma_1(1) = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \qquad w_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

On the other hand, from the relations given above for A006165(n), we can compute its linear representation:

$$v_2 = \begin{bmatrix} 1 & 1 & 1 & 0 \end{bmatrix} \qquad \gamma_2(\mathtt{0}) = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \qquad \gamma_2(\mathtt{1}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad w_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Here we are using the Iverson bracket, where (for example) the expression $[n = 0]$ evaluates to 1 if $n = 0$ and 0 otherwise.

From these two linear representations, we can easily compute the linear representation for $f(n+1) - 2 \cdot$ A006165(n) and then minimize it using the algorithm in [3, Chapter 2]. When we do so, we get a linear representation of rank 1 that evaluates to the function $-2[n = 0]$, so indeed $f(n+1) = 2 \cdot$ A006165(n) for all $n \geq 1$.

We can do the same thing for $g(n+1)$, using the `Walnut` command:

```
eval mabba n "$firstocc(i,n+1) & $abbapat(i,n+1)":
```

The resulting linear representation is

$$v_3 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \gamma_3(\mathtt{0}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \qquad \gamma_3(\mathtt{1}) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \qquad w_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

From the relations for A060973(n) given above, we can compute its linear representation:

$$v_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \qquad \gamma_4(\mathtt{0}) = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \gamma_4(\mathtt{1}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad w_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Once again we can compute the linear representation for $g(n+1) -$ A060973(n) and minimize it. When we do so, we get a linear representation of rank 0, computing the constant function 0. □

Finally, using the known expressions for the two sequences A006165 and A060973, we arrive at the following result:

Corollary 4.2. *For $n \geq 2$ we have*

$$f(n) = \begin{cases} 2^k, & \text{if } 3 \cdot 2^{k-2} < n \leq 2^k + 1; \\ 2n - 2^k - 2, & \text{if } 2^k + 1 < n \leq 3 \cdot 2^{k-1}. \end{cases}$$

For $n \geq 3$ we have

$$g(n) = \begin{cases} 2^{k-1}, & \text{if } 2^k + 1 < n \leq 3 \cdot 2^{k-1} + 1; \\ n - 2^{k-1} - 1, & \text{if } 3 \cdot 2^{k-1} + 1 < n \leq 2^k + 1. \end{cases}$$

References

[1] J.-P. Allouche and J. O. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, In *Sequences and Their Applications*, Proc. SETA '98, (Eds. C. Ding, T. Helleseth and H. Niederreiter), pp. 1–16, Springer-Verlag, 1999.

[2] C. Bernhardt, Evil twins alternate with odious twins, *Math. Mag.* **82** (2009), 57–62.

[3] J. Berstel and C. Reutenauer, Noncommutative Rational Series with Applications, Vol. 137 of *Encyclopedia of Mathematics and Its Applications*, Cambridge University Press, 2011.

[4] H.-K. Hwang, S. Janson and T.-H. Tsai, Exact and asymptotic solutions of a divide-and-conquer recurrence dividing at half: Theory and applications, *ACM Trans. Algorithms* **13** (4) (2017), #47.

[5] M. Morse, Recurrent geodesics on a surface of negative curvature, *Trans. Amer. Math. Soc.* **22** (1921), 84–100.

[6] H. Mousavi, Automatic theorem proving in `Walnut`, Arxiv preprint arXiv: 1603.06017 [cs.FL], http://arxiv.org/abs/1603.06017, (2016).

[7] J. Shallit, *The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words with* `Walnut`, London Math. Soc. Lecture Note Series, Vol. 482, Cambridge University Press, 2022.

[8] J. Shallit and R. Zarifi, Circular critical exponents for Thue-Morse factors, *RAIRO Inform. Théor. App.* **53** (2019), 37–49.

[9] N. J. A. Sloane et al., The on-line encyclopedia of integer sequences, (2022), https://oeis.org.

[10] L. Spiegelhofer, Gaps in the Thue-Morse word, preprint: arXiv:2102.01018 [math.CO], https://arxiv.org/abs/2102.01018; *J. Aust. Math. Soc.* (2022) (to appear).

[11] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1** (1912), 1–67. Reprinted in: *Selected Mathematical Papers of Axel Thue*, (Ed. T. Nagell), Universitetsforlaget, Oslo, 1977, pp. 413–478.

(Received 19 Mar 2022; revised 11 Sep 2022)