Abstract

Partial words are sequences over a finite alphabet that may contain some undefined positions called holes. We consider unavoidable sets of partial words of equal length. We compute the minimum number of holes in sets of size three over a binary alphabet (summed over all partial words in the sets). We also construct all sets that achieve this minimum. This is a step towards the difficult problem of fully characterizing all unavoidable sets of partial words of size three.

Keywords: Automata and formal languages; Combinatorics on words; Partial words; Unavoidable sets.

1 Introduction

An unavoidable set of (full) words $X$ over a finite alphabet $A$ is one for which every two-sided infinite word over $A$ has a factor in $X$ (when a word $w$ has no factor in $X$, we say that $w$ avoids $X$). For example, the set $X = \{aa, ba, bb\}$ is unavoidable over the alphabet $\{a, b\}$, since avoiding $aa$ and $bb$ forces a word to be an alternating sequence of $a$’s and $b$’s. This fundamental concept was explicitly introduced in 1983 in connection with an attempt to characterize the rational languages among the context-free ones [8]. Since then it has...
been consistently studied by researchers in both mathematics and theoretical computer science (see for example [5, 6, 7, 9, 10, 12, 13, 14]).

Partial words are sequences that may contain some undefined positions called holes, denoted by $\diamond$'s, that match every letter of the alphabet (we also say that $\diamond$ is compatible with each letter of the alphabet). For instance, $a\diamond bca\diamond b$ is a partial word with two holes over $\{a, b, c\}$, while $aabcabb$ is a full word over $\{a, b, c\}$ built by filling in the first hole with an $a$ and the second hole with a $b$. A set of partial words $X$ over $A$ is unavoidable if every two-sided infinite full word over $A$ has a factor compatible with an element in $X$.

Unavoidable sets of partial words were introduced in [3], where the problem of characterizing such sets of cardinality $n$ over a $k$-letter alphabet was initiated. Note that if $X$ is unavoidable, then every two-sided infinite unary word has a factor compatible with a member of $X$; thus $X$ cannot have fewer elements than the alphabet, and so $k \leq n$ (note that the cases $n = 1$ and $k = 1$ are trivial). The characterization of all unavoidable sets of cardinality $n = 2$ was settled recently in [2] using deep arguments related to Cayley graphs. So our next long-term goal is to characterize unavoidable sets of cardinality $n = 3$. Since in [3], all such sets over a three-letter alphabet were completely characterized (in fact, there are no non-trivial such sets), we need to focus on sets over a two-letter alphabet.

In [3], a complete characterization of all three-word unavoidable sets over a binary alphabet where each partial word has at most two defined positions was given, and some special cases where one partial word has more than two defined positions were discussed, but general criteria for these sets had not been found. In this paper, among other things, we answer affirmatively a conjecture that was left open there. Our main goal however is to make another step towards the full $n = 3$ characterization by computing the minimum number of holes in any unavoidable set of partial words of equal length and of cardinality three over a binary alphabet. We also construct all sets that achieve this minimum.

Our paper is organized as follows: In Section 2, we present the basic definitions and terminology regarding unavoidable sets of partial words that we use throughout the paper. In Section 3, we formally state our main goal towards the major problem on unavoidable sets we are concerned with, that is, the characterization problem or the problem of characterizing unavoidable sets of partial words of cardinality $n$ over a $k$-letter alphabet. In Sections 6 and 7, we make two steps towards this problem. More specifically, our first step is that we give an answer to the above mentioned conjecture on unavoidable sets of size three, while our second step is that we also compute the minimum number of holes in unavoidable sets of size three based on our characterization of these sets in two families (given in Sections 4 and 5). Finally in Section 8, we conclude with some remarks.
2 Unavoidable Sets of Partial Words

In this section, we present the basics on unavoidable sets of partial words together with the notation that we use throughout the paper. We refer the reader to Reference [1] for more background material.

Let $A$ be a fixed non-empty finite set called an alphabet whose elements we refer to as letters. A finite (full) word $w$ over $A$ is a finite sequence of letters of $A$. The sequence of length zero, or the empty word, is denoted by $\varepsilon$. We write $|w|$ to denote the length of $w$, and $w(i)$ to denote the letter at position $i$. By convention, we begin indexing the positions with 0, so a word $w$ of length $m$ can be represented as $w = w(0) \cdots w(m - 1)$. Formally, a finite word of length $m$ is a function $w : \{0, \ldots, m - 1\} \to A$. The number of occurrences of the letter $a$ in $w$ is denoted by $|w|_a$. We denote by $A^*$ the set of all finite words over $A$.

A two-sided infinite (full) word $w$ over $A$ is a function $w : \mathbb{Z} \to A$. For a positive integer $p$, $w$ is $p$-periodic or is of period $p$, if $w(i) = w(i + p)$ for all $i \in \mathbb{Z}$. We say $w$ is periodic if it has a period. If $v$ is a non-empty finite word, then $v^\mathbb{Z}$ denotes the unique two-sided infinite word $w$ with period $|v|$ such that $v = w(0) \cdots w(|v| - 1)$. Similarly, a one-sided infinite (full) word $w$ over $A$ is a function $w : \mathbb{N} \to A$. A finite word $u$ is a factor of $w$ if some integer $i$ satisfies $u = w(i) \cdots w(i + |u| - 1)$. An $m$-factor is a factor of length $m$.

A partial word $w$ of length $m$ over $A$ is a function $w : \{0, \ldots, m - 1\} \to A_\diamond$, where $A_\diamond = A \cup \{\diamond\}$ with $\diamond \notin A$. For $0 \leq i < |w|$, if $w(i) \in A$, then $i$ belongs to the domain of $w$, denoted by $D(w)$. Otherwise, $i$ is in the set of holes of $w$, denoted by $H(w)$. We denote by $A_\diamond^*$ the set of all words over $A_\diamond$ (i.e. the set of all partial words over $A$, including the empty word, $\varepsilon$). Note that full words are simply partial words without holes, that is, partial words whose domain is the entire set $\{0, \ldots, |w| - 1\}$. Two partial words $u$ and $v$ of equal length are compatible, denoted by $u \uparrow v$, if $u(i) = v(i)$ whenever $i \in D(u) \cap D(v)$. In this sense, we may view a hole as a “wildcard” character that can match any letter in $A$. We denote by $h(w)$ the number of holes in $w$, thus, $h(w) = |w| - |D(w)|$.

Let $w$ be a two-sided infinite word and let $u$ be a partial word. We say $w$ meets $u$ if $w$ has a factor compatible with $u$, and $w$ avoids $u$ otherwise. Now, $w$ meets a set of partial words $X$ if it meets some $u \in X$, and $w$ avoids $X$ otherwise. If $X$ is avoided by some two-sided infinite word, then $X$ is unavoidable; otherwise, $X$ is unavoidable or every two-sided infinite word has a factor compatible with an element in $X$. For example, the set $X = \{a, \text{bob}\}$ is unavoidable over $\{a, b\}$, since avoiding $a$ forces a word to be a sequence of $b$’s. We say $X$ is $m$-uniform if every partial word in $X$ has length $m$.

The partial word $u$ is contained in the partial word $v$, denoted by $u \subset v$, if $|u| = |v|$ and $u(i) = v(i)$, for all $i \in D(u)$. We say that $v$ is a strengthening of $u$ if $v$ has a factor containing $u$, and write $v \triangleright u$ (in other words, $v$ has a
factor built by “filling in” a number of holes in $u$). We also say that $u$ is a weakening of $v$. The following illustrates an example:

$$u = b \odot \odot \odot a$$
$$v = b \ a \ b \ \odot \ \odot \ a \ a \ b \ b \ b$$

Note that if a two-sided infinite word $w$ meets the partial word $v$, it also meets every weakening of $v$, and if $w$ avoids $u$ then $w$ avoids every strengthening of $u$.

Let $X, Y$ be sets of partial words. We extend the notions of strengthening and weakening as follows. We say that $X$ is a strengthening of $Y$ (written as $X \succ Y$) if, for each $v \in X$, there exists $u \in Y$ such that $v \succ u$. We also say that $Y$ is a weakening of $X$. For example,

$$X = \{b\odot aab, bab\odot aabbb\} \succ Y = \{b\odot a, b\odot a\odot b, aa\}$$

It is not hard to see that if the two-sided infinite word $w$ meets $X$, then it also meets every weakening of $X$, and if $w$ avoids $X$ then it avoids any strengthening of $X$. Hence if $X$ is unavoidable, so are all weakenings of $X$, and if $X$ is avoidable all strengthenings of $X$ are avoidable.

Two partial words $u$ and $v$ are conjugate, denoted by $u \sim v$, if there exist partial words $x, y$ such that $u \subset xy$ and $v \subset yx$. It is well-known that conjugacy on full words is an equivalence relation, and we use $c(m, k)$ to denote the number of conjugacy classes of words of length $m$ over a $k$-letter alphabet. However, in the case of partial words, conjugacy is no longer an equivalence relation [1]. We define two partial words $u, v$ as being hole-conjugate if there exist partial words $x, y$ such that $u = xy$ and $v = yx$; in this case we write $u \sim_\odot v$.

We conclude with some number theoretic notation used in this paper. We write $a \mid b$ if $a$ divides $b$. Next, let $p$ be a prime and let $e, m \in \mathbb{N}$. We write $p^e \parallel m$ if $p^e$ maximally divides $m$, that is, if $p^e \mid m$ but $p^{e+1} \nmid m$. Finally, we write $i \equiv_m j$ if $i$ is congruent to $j$ modulo $m$.

## 3 The Characterization Problem on Unavoidable Sets

In this paper, we are concerned with the characterization problem, that is, the problem of characterizing unavoidable sets of partial words of cardinality $n$ over a $k$-letter alphabet. We make two steps towards this problem. As a first step, we answer affirmatively a conjecture by Blanchet-Sadri et al. regarding the maximum number of interior defined positions in unavoidable sets of the form $\{a\odot a^{m-2}a, b\odot a^{m-2}b, x\}$ where $x$ is compatible with $b\odot a^{m-2}a$ (Conjecture 2 of [3]). As a second step, as we are interested in unavoidable sets with the minimum number of holes, and strengthenings do not contain more holes than the original set, it is reasonable to investigate “maximal
strength” unavoidable sets. So let $X$ be an unavoidable set. If, for all $Y > X$, $Y$ is avoidable, then we say $X$ is maximal. We calculate the minimum number of holes in any unavoidable $m$-uniform set (summed over all partial words in the set) of cardinality three over a binary alphabet. We construct all sets that achieve this minimum, and then show that any unavoidable set with the stated number of holes is maximal.

As discussed earlier, we can restrict our attention to the binary alphabet $\{a, b\}$. Hence, we may refer to $a$ and $b$ as complements of each other, so that $\overline{a} = b$ and $\overline{b} = a$. A two-sided infinite word $w$ is $p$-alternating if $w(i) = \overline{w(i+p)}$ for all $i \in \mathbb{Z}$. Note that if $w$ is $p$-alternating, it is also $2p$-periodic.

We denote by $H_{m,n}$ the minimum number of holes in any unavoidable $m$-uniform set (summed over all partial words in the set) of cardinality $n$ over a binary alphabet. To have words of “real length” $m$, we require that $D(u) \ni 0, m-1$ for each $u$ in any such set. The minimum number of elements in an unavoidable set of full words of length $m$ over $\{a, b\}$ is known to be equal to the number $c(m,2)$ of conjugacy classes of words of length $m$ over $\{a, b\}$ [11, 5]. Thus, $H_{m,c(m,2)} = 0$ for $m \geq 1$.

**Proposition 1.** If every $m$-uniform unavoidable set of cardinality $n$ having a total of $h$ holes is maximal, then $H_{m,n} \geq h$.

**Proof.** If $h = 0$ then the claim is clear, so assume $h \geq 1$. Suppose that $H_{m,n} < h$, and let $Y$ be an $m$-uniform unavoidable set of cardinality $n$ with $h' < h$ holes for some $h' \in \mathbb{N}$. Now add holes to words in $Y$ arbitrarily until the new set, $Y'$, has $h$ holes. Since $Y' \prec Y$, $Y'$ is also unavoidable. Hence $Y'$ is an $m$-uniform unavoidable set that is not maximal. $\square$

We now state the main result and focus of this paper.

**Theorem 1.** For $m \geq 4$, $H_{m,3} = 2m - 5$ if $m$ is even, and $H_{m,3} = 2m - 6$ if $m$ is odd.

**Remark 1.** As long as we are discussing an $m$-uniform unavoidable set of size three, say $X = \{x_1, x_2, x_3\}$, we may always assume, without loss of generality:

- $x_1(0) = x_2(m-1) = a$, and only $a$’s and $\diamond$’s appear in $x_1$;
- $x_2(0) = x_2(m-1) = b$, and only $b$’s and $\diamond$’s appear in $x_2$;
- $x_3(0) = b$ and $x_3(m-1) = a$;
- $h(x_1) \leq h(x_2)$.

We call this the standard form of an $m$-uniform three-element unavoidable set of partial words. The presence of $x_1, x_2$ is justified since any unavoidable set over $\{a, b\}$ must contain words compatible with $a^\mathbb{Z}$ and $b^\mathbb{Z}$,
respectively. Now, $x_3$ must have complementary ends, since otherwise $X = \{a^{m-2}a, b^{m-2}b\}$ and as the latter set is avoidable so is $X$. Next, if $h(x_1) > h(x_2)$, we may consider instead the set $\{x_1, x_2, x_3\}$. This “switches” the identity of $x_1$ and $x_2$ so that $h(x_1) \leq h(x_2)$). Finally, we may fix the orientation of $x_3$ by taking the reverse of each word, if necessary.

In the next two sections, we give constructions of sets that achieve the proposed minimum of Theorem 1.

4 The $C$-Sets

In this section, we define and completely characterize the unavoidable $C$-sets.

**Definition 1.** Let $\Lambda \subset \{1, \ldots, m-2\}$. We denote by $C_m(\Lambda)$ the $m$-uniform set $\{x_1, x_2, x_3\}$ where $x_1 = a^m$, $x_2 = b^{m-2}b$, and $x_3$ is defined as follows:

$$x_3(i) = \begin{cases} 
  b & \text{if } i = 0, \\
  a & \text{if } i \in \Lambda \cup \{m-1\}, \\
  \circ & \text{otherwise}.
\end{cases}$$

**Remark 2.** If $\Lambda = \{i_1, i_2, \ldots, i_s\}$, we often write $C_m(i_1, i_2, \ldots, i_s)$ instead of $C_m(\{i_1, i_2, \ldots, i_s\})$. By convention, we order the arguments of $C_m(i_1, i_2, \ldots, i_s)$ in increasing order, so that $i_1 < i_2 < \cdots < i_s$.

**Remark 3.** We have $C_m(\Lambda) \prec C_m(\Gamma)$ precisely when $\Lambda \subset \Gamma$.

For the characterization of the unavoidable $C$-sets, we start with one position filled in.

**Proposition 2.** The set $C_m(i)$ is unavoidable if and only if $i \mid m-1$.

**Proof.** Suppose $i \mid m-1$ with $li = m-1$ for some $l \in \mathbb{N}$, and suppose to the contrary that $w$ is a two-sided infinite word that avoids $X = C_m(i)$. The word $w$ must contain a $b$ in order to avoid $x_1$; say, without loss of generality, that $w(0) = b$. To avoid $x_2$, it must be that $w(m-1) = a$. This, however, forces $w(i) = b$, or else $w$ meets $x_3$. We may repeat the argument to conclude that $w(l'i) = b$ for all $l' \in \mathbb{N}$. This yields a contradiction, as we claimed that $w(li) = w(m-1) = a$. Conversely, if $i \nmid m-1$, then let $w = (ba^{l-1})^2$. Now, $w$ clearly avoids $x_1$ and $x_3$ as it is $i$-periodic. Finally, all indices containing $b$ are congruent to each other modulo $i$. Thus, $w$ does not meet $x_2$, since any two positions $m-1$ apart are not congruent modulo $i$, and so cannot both be $b$. Hence, $X$ is avoidable.

Next, for two positions filled in, we have the following result.
**Proposition 3.** The set $C_m(i, j)$ is unavoidable if and only if $i, j \mid m - 1$ and $j = 2i$.

**Proof.** Suppose $i, j \mid m - 1$ with $li = m - 1$ for some $l \in \mathbb{N}$ and $2i = j$, and suppose to the contrary that $w$ is a two-sided infinite word that avoids $X = C_m(i, j)$. Note that every $b$ in $w$ must be followed by an $a$ after $m - 1$ positions (to avoid $x_2$), and be followed by a $b$ after either $i$ or $j$ positions (to avoid $x_3$). It is impossible that every consecutive pair of $b$'s be separated by $j$ positions, for if so $w$ meets $x_2$ (as $j \mid m - 1$). Hence, some pair of $b$'s are separated by $i$ positions; say $w(0) = w(i) = b$. This implies that $w(m - 1) = w(m - 1 + i) = a$. Now, if $w(m - 1 - i) = b$, then $w$ meets $x_3$ (since that $b$ has $a$'s both $i$ and $2i = j$ positions later). This argument cascades backwards since we once again have $a$'s separated by $i$ positions. Thus $w(m - 1 - l'i) = a$ for all $l' \in \mathbb{N}$, but this is a contradiction since $w(m - 1 - li) = w(0) = b$. Hence no word $w$ avoids $X$.

On the other hand, if $i \nmid m - 1$ then $C_m(i, j) \supset C_m(i)$, where the latter set is avoidable by Proposition 2, and so $C_m(i, j)$ is also avoidable (similarly, for the case when $j \nmid m - 1$). Finally, if $2i \neq j$ and $i, j \mid m - 1$, put $lj = m - 1$ for some $l \in \mathbb{N}$. Let $u = ba^{i-1}(ba^{j-1})^{l-1}$. Then we claim $w = u^Z$ is a two-sided infinite word avoiding $X$. Clearly $w$ avoids $x_1$ and $x_3$ (for every $b$ is followed by another one after either $i$ or $j$ positions). Now let $v$ be any $m$-factor of $w$ with $v(0) = b$. We claim that $v(m - 1) = a$ and so $w$ avoids $x_2$. Note that $b$'s appear in positions congruent to 0 modulo $j$ until the first factor of $ba^{i-1}$ appears, after which they appear in positions congruent to $i$ modulo $j$. The next time a factor of $ba^{i-1}$ appears, $b$'s start appearing in indices congruent to $2i$ modulo $j$, and so on.

Now, recall that $i < j$, and so $m = lj + 1 > lj + i - j = (l - 1)j + i = |u|$. Furthermore, since $j < m - 1$, we know that $l \geq 2$. It follows that

$$m < m - 1 + 2i \leq m - 1 + 2i + (l - 2)j = lj + 2i + lj - 2j = 2((l - 1)j + i) = 2|u|$$

Therefore, any $m$-factor $v$ of $w$ contains more than one but less than two full copies of $u$. Hence there are either one or two occurrences of $ba^{i-1}$ (which appear once per $u$). So $b$'s appear at the end of $v$ in positions congruent to $i$ or $2i$ modulo $j$. Now, the only way for $v(m - 1) = b$ is if $m - 1 \equiv j$ or $m - 1 \equiv 2i$. But $j \mid m - 1$, so $m - 1 \equiv j 0$. It is easy to see that $i \equiv 0$ is impossible since $i < j$, and $2i \equiv 0$ implies $2i = lj$ for some $l$. As $i < j$, this forces $l = 1$ and so $2i = j$, contrary to hypothesis. Hence if $v$ is an $m$-factor of $w$ with $v(0) = b$, then $v(m - 1) = a$. So, $w$ avoids $x_2$ and hence the set $X$. \[\square\]

Finally, for at least three positions filled in, we get the following as a corollary.

**Corollary 1.** If $\Lambda \subset \{1, \ldots, m - 2\}$ with $|\Lambda| \geq 3$, then $C_m(\Lambda)$ is avoidable.
Proof. Put $\Lambda = \{i_1, \ldots, i_s\}$ with $s \geq 3$. Now, $C_m(\Lambda) \supset C_m(i_1, i_2)$ and $C_m(\Lambda) \supset C_m(i_1, i_3)$, and since $i_2 \neq i_3$ at least one of $C_m(i_1, i_2)$ and $C_m(i_1, i_3)$ is avoidable by Proposition 3. Hence, so is the set $C_m(\Lambda)$. \qed

5 The $D$-Sets

In this section, we define and completely characterize the unavoidable $D$-sets.

Definition 2. Let $\Lambda \subset \{1, \ldots, m-2\}$. We denote by $D_m(\Lambda)$ the $m$-uniform set $\{x_1, x_2, x_3\}$ where $x_1 = a\omega_m^{-2}a$, $x_2 = b\omega_m^{-2}b$, and $x_3$ is defined as follows:

\[
x_3(i) = \begin{cases} 
  b & \text{if } i = 0, \\
  a & \text{if } i \in \Lambda \cup \{m-1\}, \\
  \diamond & \text{otherwise.}
\end{cases}
\]

As before, if $\Lambda = \{i_1, i_2, \ldots, i_s\}$, we often write $D_m(i_1, i_2, \ldots, i_s)$ instead of $D_m(\{i_1, i_2, \ldots, i_s\})$, and we order the arguments of $D_m(i_1, i_2, \ldots, i_s)$ in increasing order, so that $i_1 < i_2 < \cdots < i_s$.

We now characterize the unavoidable $D$-sets with one position filled in. However, this process is much more difficult than the corresponding task for $C$-sets, owing to the stricter requirements imposed by $x_1$.

Lemma 1 ([3]). Let $X = \{a\omega m^{-2}a, b\omega n^{-2}b\}$. Set $2^s \parallel m + 1$ and $2^t \parallel n + 1$. Then $X$ is unavoidable if and only if $s \neq t$.

Lemma 2. The sets $X = \{a\omega_m^{-2}a, b\omega_n^{-2}b\}$, $Y = \{a\omega_m^{-2}a, b\omega_n^{-2}b, a\omega^n^{-2}a\}$ have the same avoidability.

Proof. Suppose $X$ is avoidable, say by the two-sided infinite word $w$. Suppose that $w$ meets $a\omega_m^{-2}a$, so that $w(i) = w(i + n - 1) = a$ for some $i \in \mathbb{Z}$.

Then $w(i + m - 1) = w(i + n - 1 + m - 1) = b$, since $w$ avoids $a\omega_m^{-2}a$, but this contradicts the fact that $w$ avoids $b\omega_n^{-2}b$. Hence $w$ avoids $a\omega^n^{-2}a$ and so avoids $Y$. But clearly $X > Y$, and so if $X$ is unavoidable so is $Y$. \qed

Proposition 4. If $2^s \parallel m - 1$ and $2^t \parallel i$, then $D_m(i)$ is unavoidable if and only if $t \leq s$.

Proof. Let $X = \{b\omega_n^{-2}b, a\omega_m^{-2-i}a\}$. We first show that $X$ has the same avoidability as $D_m(i)$. For suppose $X$ is avoidable. Then so is $Y = X \cup \{a\omega_m^{-2}a\}$, by Lemma 2. As $Y$ is an unavoidable weakening of $D_m(i)$, we conclude that $D_m(i)$ is avoidable. On the other hand, suppose $X$ is unavoidable. Let $w$ be any two-sided infinite word. If $w$ meets $b\omega_m^{-2}b$, then it also meets $D_m(i)$. If it does not, then $w(j) = w(j - m + 1 + i) = a$ for some $j \in \mathbb{Z}$.  

8
Now, if \(w(j - m + 1) = a\), then \(w\) meets \(x_1\), and if \(w(j - m + 1) = b\), it meets \(x_3\). In either case, \(w\) meets \(D_m(i)\), and so \(D_m(i)\) is unavoidable. Hence \(X\) has the same avoidability as \(D_m(i)\).

Next, let \(2^s \parallel m - 1, 2^t \parallel i, 2^u \parallel m - 1 - i\). We show that \(r \neq s\) if and only if \(t \leq s\). Set \(2^p = m - 1, 2^q = i\) for odd \(p, q\). Now, if \(t < s\), then \(2^{s-t}p - q\) is odd, and so \(2^t \parallel 2^t(2^{s-t}p - q) = 2^s p - 2^t q = m - 1 - i\) and \(r = t \neq s\). If \(t = s\), then, since \(p - q\) is even, we have \(2^{s+1} | 2^s (p - q) = 2^s p - 2^s q = m - 1 - i\). Thus \(r \geq s + 1\) and so \(r\) cannot be equal to \(s\). Finally, if \(t > s\), then \(p - 2^{s-t}q\) is odd. It follows that \(2^s \parallel 2^s (p - 2^{s-t}q) = 2^s p - 2^s q = m - 1 - i\) and so \(r = s\). Hence \(r \neq s\) if and only if \(t \leq s\). Recall that by Lemma 1, \(X\) is unavoidable if and only if \(r \neq s\). Therefore, \(D_m(i)\) is unavoidable if and only if \(t \leq s\).

We now turn our attention to \(D\)-sets with two positions filled in. A previous result gives necessary conditions for the unavoidability of \(D_m(i, j)\), provided that \(i, j, m - 1\) are relatively prime.

**Theorem 2 ([2]).** Let \(l, n_1, n_2\) be non-negative integers such that \(n_1 \leq n_2\) and \(\gcd(l + 1, n_1 + 1, n_2 + 1) = 1\). If the set \(\{a^{n_1}b, b^{n_1}a, a^{n_2}b, b^{n_2}a\}\) is unavoidable, then at least one of the following conditions hold:

(i) \(l = 6\) and \((n_1, n_2) \in \{(1, 3), (3, 7), (1, 7)\};
(ii) \(n_1 + 1 \equiv 2^{l+2}\ 0;\)
(iii) \(n_2 + 1 \equiv 2^{l+2}\ 0;\)
(iv) \(n_1 + n_2 + 2 \equiv 2^{l+2}\ 0;\)
(v) \(2n_1 + n_2 + 3 \equiv 2^{l+2}\ l + 1;\)
(vi) \(2n_2 + n_1 + 3 \equiv 2^{l+2}\ l + 1;\)
(vii) \(n_2 - n_1 \equiv 2^{l+2}\ l + 1.\)

**Corollary 2.** If \(D_m(i, j)\) is unavoidable and \(\gcd(m - 1, i, j) = 1\), then \(j = 2i\), or \(i + j = m - 1\), or the three conditions \(m = 8, i = 1,\) and \(j \in \{3, 5\}\) hold.

**Proof.** Suppose \(D_m(i, j)\) is unavoidable. Put \(l = m - 2, n_1 = j - i - 1, n_2 = m - j - 2\) and let \(Y = \{a^{n_1}b, b^{n_1}a, a^{n_2}b, b^{n_2}a\}\). Note that \(Y\) is also unavoidable since \(Y \subset D_m(i, j) = \{a^{n_1}b, b^{n_1}a, b^{n_1}a, a^{n_2}b\}\); moreover, \(\gcd(l+1, n_1+1, n_2+1) = 1\). Hence, \(l, n_1, n_2\) must satisfy one of the conditions given in Theorem 2. However, as \(i > 0\) we have that \(n_1 + n_2 + 1 < l\); this forces one of (i), (v), or (vi) to hold. It is easy to verify that these conditions are equivalent to the ones stated about \(m, i, j\).

The following proposition shows that we do not gain any new unavoidable sets by considering cases where \(m - 1, i, j\) are not relatively prime. Thus we may extend the above result to all \(i, j, m\).
Proposition 5. For any $\Lambda = \{i_1, \ldots, i_s\}$, let $d\Lambda = \{di \mid i \in \Lambda\}$.
Then $D_m(\Lambda)$ is avoidable if and only if $D_{d(m-1)+1}(d\Lambda)$ is.

Proof. Let $\Lambda = \{i_1, \ldots, i_s\} \subset \{1, \ldots, m-2\}$. Let $Y = D_m(\Lambda) = \{y_1, y_2, y_3\}$ and $Z = D_{d(m-1)+1}(d\Lambda) = \{z_1, z_2, z_3\}$, where $y_1 = a^{s-m-2}a, y_2 = b^{s-m-2}b, z_1 = a^{d(m-1)-1}a, z_2 = b^{d(m-1)-1}b$. If $w$ is a word avoiding $Y$, then we claim the word $w' = \cdots w(-1)^{d}w(0)^{d}w(1)^{d}\cdots$ avoids $Z$. To see this, note that as $w$ is $(m-1)$-alternating, $w'$ is $d(m-1)$-alternating and so avoids $z_1, z_2$. Now, if $w'$ meets $z_3$, then there exists $l$ such that $w'(l) = b, w'(l + d_i) = \cdots = w(l + d_i - 1) = w(l + d(m-1)) = a$. But if we put $h = \lfloor \frac{l}{d} \rfloor$, then $w(h) = b, w(h + i) = \cdots = w(h + i_s) = w(h + m - 1) = a$ so $w$ meets $y_3$. This is a contradiction, so $w'$ in fact avoids $z_3$ and hence $Z$. The reverse direction is analogous, except that if $w$ is a word avoiding $Z$, then the word $w' = \cdots w(-d)w(0)w(d)\cdots$ avoids $Y$.

Corollary 3. If $D_m(i, j)$ is unavoidable, then $j = 2i$, or $i + j = m - 1$, or both $m = 7i + 1$ and $j \in \{3i, 5i\}$.

Proof. This is an immediate consequence of Corollary 2 and Proposition 5. □

We now show that the above conditions are sufficient.

Lemma 3. Let $m, n \in \mathbb{N}$, $2^s \parallel m$ and $2^t \parallel n$. If $s \geq t$, $\gcd(m, n) = \gcd(2m, n)$.

Proof. Since $s \geq t$, we know that the power of 2 maximally dividing $\gcd(m, n)$ is just $\min(s, t) = t$. But the power of 2 maximally dividing $\gcd(2m, n)$ is $\min(s + 1, t) = t$. It is clear that the other prime factors of $\gcd(m, n)$ are unaffected by doubling $m$, and the result follows. □

Proposition 6. Let $2^s \parallel m - 1, 2^t \parallel i$, and $2^r \parallel j$. Then the set $D_m(i, j)$ is unavoidable if and only if (iv) holds in addition to one of (i), (ii), or (iii):

(i) $j = 2i$;
(ii) $i + j = m - 1$;
(iii) $m = 7i + 1$ and $j \in \{3i, 5i\}$;
(iv) $s \geq t, r$.

Proof. If $t > s$, then $D_m(i)$ is avoidable by Proposition 4. Hence $D_m(i, j)$ is avoidable, as $D_m(i, j) \succ D_m(i)$. A similar argument applies if $r > s$. Together with Corollary 3, we have one direction of the proof.

It remains to show that the above conditions are sufficient. We assume for the remainder of the proof that (iv) holds.
Suppose (i) holds, and that \( w \) is a word avoiding \( D_m(i, j) \). We show that this leads to a contradiction. Since \( w \) avoids \( x_1 \), we have \( |w|_b \geq 1 \) and we may take without loss of generality \( w(0) = b \). To avoid \( x_2 \), \( w(m - 1) = a \), and to avoid \( x_3 \), \( w(i) = b \) or \( w(j) = b \). Similarly, for every \( b \), there must be a \( b \) that occurs \( i \) or \( j = 2i \) positions later. Suppose that \( w(i) = b \). Then \( w(m - 1 + i) = a \). Now, note that \( w(m - 1 - i) = a \), for there are \( a \)'s that occur \( i \) positions and \( j = 2i \) positions after \( m - 1 - i \). Thus \( w(-i) = b \). Since we have another two \( a \)'s separated by \( i \) positions (at \( m - 1 \) and \( m - 1 - i \)), we may apply the same argument to conclude that \( w(-2i) = b \). We may repeat this to get \( w(li) = b \) for all \( l \leq 0 \). Now, \( w \) is \((m-1)\)-alternating since it avoids \( \{x_1, x_2\} \), and so it is \((2m-2)\)-periodic. Hence \( w(x) = b \) whenever \( x \equiv_{2m-2} li \) for some \( l \leq 0 \).

Let \( d = \gcd(m - 1, i) \). Then \( d \mid m - 1 \), say with \( dq = m - 1 \), and furthermore \( d = \gcd(2m - 2, i) \) by Lemma 3. By Bezout’s theorem, we may write \( d = xi + y(2m - 2) \) for some \( x, y \in \mathbb{Z} \) (\( x \) negative). Hence \( xi \equiv_{2m-2} d \). It follows that \( w(m - 1) = w(dq) = b \), as \( dq \equiv_{2m-2} xqi \). This contradicts our previous assertion that \( w(m - 1) = a \).

It remains to consider the case where \( b \) appears in every position congruent to \( lj \) modulo \( 2m - 2 \) for some \( l \in \mathbb{Z} \) (that is, when no two \( b \)'s are separated by \( i \) positions), but this leads to a contradiction in the same way, since \( r \leq s \). Hence we may represent \( m - 1 \) as a multiple of \( j \) modulo \( 2m - 2 \) and so reach a contradiction. We conclude that \( D_m(i, j) \) is unavoidable when (i) holds.

Now suppose (ii) holds. Again, let \( w \) be a word that avoids \( D_m(i, j) \), and take without loss of generality \( w(0) = b \). Suppose that \( w(i) = b \). Then \( w(m - 1) = w(m - 1 + i) = a \). Now, the \( b \) in position \( i \) already has an \( a \) \( m - 1 - i = j \) positions later, so it must have a \( b \) \( i \) positions later. Hence \( w(2i) = b \), and now \( w(m - 1 + 2i) = a \). Repeating this argument gives us that \( w(li) = b \) for all \( l \geq 0 \). Since \( w \) is \((2m-2)\)-periodic, we have \( w(x) = b \) whenever \( x \equiv_{2m-2} li \) for some \( l \). A contradiction is obtained in a manner identical to the previous case, since (iv) holds. Hence \( D_m(i, j) \) is unavoidable when (ii) holds. Finally, note that there are only a finite number of words that are \((m-1)\)-alternating, for any fixed \( m \). Thus we may show the unavoidability of \( D_8(1, 3) \) and \( D_8(1, 5) \) (and hence the unavoidability of \( D_{7i+1}(i, 3i) \) and \( D_{7i+1}(i, 5i) \), by Proposition 5) via an exhaustive search. It follows that \( D_m(i, j) \) is unavoidable if (iii) holds.

Finally, we show that, like the \( C \)-sets, the \( D \)-sets are always avoidable when \( x_3 \) has at least three positions filled in.

**Proposition 7.** If \( \Lambda \subset \{1, \ldots, m-2\} \) with \( |\Lambda| \geq 3 \), then \( D_m(\Lambda) \) is avoidable.

**Proof.** It suffices to show that \( D_m(i, j, l) \) is avoidable, as if \( |\Lambda| > 3 \) we can choose a weakening with exactly three positions filled in \( x_3 \). Moreover, by Proposition 5, we only need to consider the cases when \( \gcd(m - 1, i, j, l) = 1 \).
If $D_m(i, j, l)$ is unavoidable, then it is necessary that each of the sets $D_m(i, j)$, $D_m(j, l)$, and $D_m(i, l)$ be unavoidable. Hence each weakening must satisfy Proposition 6. Suppose some of these three weakenings satisfies (iii). If $m = 8$ it is easy to see that one of the above weakenings of $D_m(i, j, l)$ is avoidable, as $D_8(1, 3)$ and $D_8(1, 5)$ are the only unavoidable $D$-sets. On the other hand, suppose $m = 7d + 1$ with $d > 1$. If $D_m(i, j)$ satisfies (iii), then $l$ is also a multiple of $d$ regardless of which condition $D_m(i, l)$ satisfies. This contradicts our claim of relative primeness. An analogous argument shows that $D_m(i, l)$ cannot satisfy (iii).

Now suppose $D_m(j, l)$ satisfies (iii). Then $j = d$ and $l = pd$ for $p \in \{3, 5\}$. If $D_m(i, j)$ satisfies (ii) then again $i$ is a multiple of $d$ and we have a contradiction. Hence $D_m(i, j)$ satisfies (i) and $j = 2i$. If $i > 1$ we again contradict relative primeness (since $\gcd(m - 1, i, j, l) = i$), and if $i = 1$, we have $d = 2$. But both $D_{15}(1, 6), D_{15}(1, 10)$ are avoidable, so $D_m(i, j, l)$ has the avoidable weakening $D_m(i, l)$. Hence if any of the three weakenings satisfy (iii), $D_m(i, j, l)$ is avoidable.

Next suppose none of the three weakenings satisfies (iii). Set $2^s \parallel m - 1, 2^r \parallel i, 2^t \parallel l$. It is impossible that all three weakenings satisfy (i), just as it is impossible for more than one weakening to satisfy (ii). Hence it must be that two weakenings satisfy (i) and one weakening satisfies (ii). It is easy to see that we must have $j = 2i, l = 2j$, and $i + l = m - 1$. But this implies $l = 4i$, and so $5i = m - 1$. It follows that $s = t$. Hence we have $r > s$, which is a contradiction as we assumed (iv) holds. Therefore, $D_m(i, j, l)$ is avoidable.

With our characterization of unavoidable $C$-sets and $D$-sets, we may begin to prove Theorem 1. We first prove Conjecture 2 from [3].

6 Answer to a Conjecture on Unavoidable Sets of Size Three

Corollary 4 answers the following conjecture.

**Conjecture 1 ([3]).** If the set $X = \{a \diamond^{m-2} a, b \diamond^{m-2} b, x\}$ is unavoidable, where $x \uparrow b \diamond^{m-2} a$, then $x$ has at most two interior defined positions.

We begin with a lemma.

**Lemma 4.** Let $i_1 < \cdots < i_s < j_1 < \cdots < j_r$ be elements of $\{1, \ldots, m - 2\}$. Let $x$ be defined as follows: $x(i) = b$ if $i \in \{0, i_1, \ldots, i_s\}$, $x(i) = a$ if $i \in \{j_1, \ldots, j_r, m - 1\}$, and $x(i) = \diamond$ otherwise. Then the set $X = \{a \diamond^{m-2} a, b \diamond^{m-2} b, x\}$ has the same avoidability as some $D$-set $D_m(\Lambda)$ with $|\Lambda| = s + r$. 
Proof. We proceed by induction on $s$. The base case of $s = 0$ is trivial as then $X$ is itself a $D$-set. Now let $s \geq 1$. Note that a word $w$ meets $x$ if and only if it meets $x'$ defined as

$$b_{i_2}^{j_2 - i_1} \cdots b_{i_s}^{j_s - i_{s-1}} a_{i_1}^{m-1-j_s} \cdots a_{i_s}^{m-1-j_1} a$$

since $w$ must be $(m-1)$-alternating. Hence $X$ has the same avoidability as $X' = \{a_{m-2}^o a, b_{m-2}^o b, x'\}$ which has one fewer $b$. Applying the induction hypothesis to $X'$ yields the claim.

Corollary 4. Conjecture 1 is true.

Proof. If $x$ has any $a$ appearing before a $b$, then the set $X$ is avoided by $(b_{m-1}^o a_{m-1}^o \mathbb{Z})$. Otherwise, if $x$ has at least three interior defined positions, then by Lemma 4 it has the same avoidability as some set $D_m(\Lambda)$ with $|\Lambda| \geq 3$. But all such $D$-sets are avoidable, by Proposition 7, and so $X$ is avoidable.

7 Minimum Number of Holes in Unavoidable Sets of Size Three

First, we show that the $C$-sets are the only unavoidable sets with the minimum number of holes. We divide the sets into multiple cases, conditioning on the quantity $h(x_1) + h(x_2)$.

Corollary 5. Let $m$ be odd (resp., even). Let $X$ be an $m$-uniform set of size three of the form described in Remark 1. Suppose $h(x_1) + h(x_2) > m - 2$ (resp., $m - 1$). Then if $X$ has $2m - 6$ (resp., $2m - 5$) holes in total, $X$ is avoidable.

Proof. There are at most $m - 5$ holes in $x_3$, and so $x_3$ has at least three positions other than 0 and $m - 1$ defined. Then we may weaken $x_1, x_2$ to $a_{m-2}^o a, b_{m-2}^o b$. The resulting set is avoidable by Corollary 4, and therefore so is $X$.

Note that we did not treat the case where $h(x_1) + h(x_2) = m - 1$ for even $m$. This case is covered by the following proposition.

Proposition 8. Let $m \geq 4$ be even, and let $X$ be an $m$-uniform set of size three of the form described in Remark 1 with $h(x_1) + h(x_2) = m - 1$. Then if $X$ has $2m - 5$ holes in total, $X$ is avoidable.

Proof. First, suppose that $h(x_1) > 1$. Assume that $m \geq 8$. We find a two-sided infinite word $w$ with period $m - 1$ that avoids $X$. Since $w$ is $(m - 1)$-periodic, any $m$-factor of $w$ begins and ends with the same letter, and so $w$ immediately avoids $x_3$. Moreover, we only have to consider whether
and its period, respectively). Unless either

\( w \) meets \( x_1' = x_1(0) \cdots x_1(m - 2) \) (and \( x_2' = x_2(0) \cdots x_2(m - 2) \)), as any \( m \)-factor \( v \) with \( v(0) = a \) necessarily has \( v(m - 1) = a \) (analogously, every \( m \)-factor that begins with \( b \) has to end with \( b \)).

Now consider the set \( B \), which contains all conjugacy classes of length \( m - 1 \) over \( \{a, b\} \), with exactly \( h(x_1) \) \( b \)'s and \( h(x_2) \) \( a \)'s. Since \( m \geq 8 \), it follows that \( |B| > 2 \). Choose a representative \( u \) of a conjugacy class not covered by \( x_1' \) and \( x_2' \). By considering the number of \( a \)'s and \( b \)'s in \( u \), we see that if \( w = u^x \) were to meet \( x_1' \) via the \((m - 1)\)-factor \( v \), the \( \phi \)'s in \( x_1' \) need to align with the \( b \)'s in \( v \). However, for any factor \( v \) of \( w \) this is impossible, since \( u \sim x_1' \) and \( v \sim u \). Thus, it follows that \( v \) cannot be compatible with \( x_1' \). A similar argument shows that \( w \) avoids \( x_2' \). Hence \( w \) avoids \( x_1 \) and \( x_2 \), and therefore avoids \( X \). We may check the cases for \( m \leq 6 \) easily via a computer program.

Now, suppose that \( h(x_1) = 1 \). In this case we know that \( x_1 \sim_\phi a^{m-1}b \) and \( x_2 = b \omega^{m-2}b \). Moreover, \( x_3 \) has precisely two interior positions defined. First, if both the interior positions have letter \( b \), then the word \( w_1 = (baba^{m-3})^Z \) avoids \( X \) since each \( m \)-factor of \( w_1 \) contains exactly two occurrences of the letter \( b \), and so cannot be compatible with either \( x_1 \) or \( x_3 \). The word \( w_1 \) avoids \( x_3 \) as well since both \( m \)-factors that begin with \( b \) end with \( a \). Second, if the interior positions have letters, from left to right, \( a, b, b \), then the word \((b^{m-1}a^{m-1})^Z \) avoids \( X \). Third, if the interior positions have letters, from left to right, \( b, a, b \), and the \( b \) occurs in position 1, then \((baba^{m-3})^Z \) avoids \( X \). Otherwise, the word \((bba^{m-1})^Z \) avoids \( X \), since in any \( m \)-factor which contains two instances of \( b \), these letters appear in consecutive positions, and so cannot be compatible with \( x_2 \) or \( x_3 \).

Finally, if both the interior positions \( i, j \), and \( i < j \), have letter \( a \), then we proceed as follows. If \( i, j \mid m - 1 \), then, since \( m - 1 \) is odd it cannot be that \( j = 2i \). Therefore the word \( w_2 = (ba^{l-1}(ba^{j-1})^{l-1})^Z \) (where \( jl = m - 1 \)) avoids the set \( C_m(i, j) \) by Proposition 3, and so avoids \( x_2 \) and \( x_3 \). Since \( w_2 \) has at least two occurrences of \( b \) in each \( m \)-factor, \( w_2 \) avoids \( x_1 \) as well. Hence \( w_2 \) avoids \( X \).

If \( i \) and \( j \) do not simultaneously divide \( m - 1 \), let \( l \in \{i, j\} \) be an index that does not divide \( m - 1 \). Now, \((ba^{l-1})^Z \) avoids \( x_2 \) and \( x_3 \), but it might meet \( x_1 \) if the number of \( a \)'s on either side of the \( \phi \) in \( x_1 \) are both less than \( l \). This can happen only if \( l > \frac{m}{2} \), which in turn implies that \( j > \frac{m}{2} \) (either \( l = j \) or \( l = i \)). Hence \( j \mid m - 1 \) as well. Then the \( j \)-periodic word \( w_3 = (bba^{j-2})^Z \) avoids \( x_1 \) and \( x_3 \) (consider the number of instances of \( b \) in \( w_3 \) and its period, respectively). Unless either \( j + 1 = m - 1 \) or \( 2j - 1 = m - 1 \), the word \( w_3 \) avoids \( x_2 \) as well. However, in both of these last cases the word \((bba^{j-3})^Z \) avoids \( X \).

\[ \square \]

**Proposition 9.** Let \( X \) be an \( m \)-uniform set of three partial words of the form described in Remark 1. If \( h(x_1) + h(x_2) = m - 2 \), then either \( X \) is a \( C \)-set or \( X \) is avoidable.
Proof. Suppose \( h(x_1) = 0 \). Then if \( |x_3|_b \geq 2 \), the two-sided infinite word \( w = (ba^{m-1})^\mathbb{Z} \) avoids \( X \); otherwise, \( X \) is a \( C \)-set. Therefore, for the remainder of this proof we may assume that \( h(x_1) \geq 1 \). For brevity, let \( h(x_1) = i - 2 \). Then \( h(x_2) = m - i \).

First, suppose that \( x_2 \sim \ (b^i a^{m-i})^\mathbb{Z} \). The word \( w = (b^{i-1} a^{m-i})^\mathbb{Z} \) avoids \( X \). Note that \( w \) is \((m - 1)\)-periodic, so \( w \) does not meet \( x_3 \) (any \( m \)-factor of \( w \) has the same symbol in its first and last position). Since any \( m \)-factor of \( w \) has at least \( i - 1 \) \( b \)'s, while \( x_1 \) contains only \( i - 2 \) \( \circ \)'s, we can conclude that \( w \) avoids \( x_1 \). Finally, let \( v \) be any \( m \)-factor of \( w \) with \( v(0) = b \). Then \( v(m - 1) = b \) as \( w \) is \((m - 1)\)-periodic, and \( v(0) \cdots v(m - 2) \sim b^{i-1} a^{m-i} \).

This implies that there exists a contiguous block of \( m - i \) \( a \)'s within \( v \). It is now clear that \( v \not\sim x_2 \), as \( x_2 \) has precisely \( m - i \) \( \circ \)'s to match the \( a \)'s, but they do not form a contiguous block. By assumption \( v \) is any \( m \)-factor of \( w \) that begins with a \( b \), we can therefore conclude that \( w \) avoids \( x_2 \) and hence the set \( X \).

Now, suppose that \( x_2 \sim \ (b^i a^{m-i})^\mathbb{Z} \). The word \( w_1 = (b^{i-2} aba^{m-i-1})^\mathbb{Z} \) avoids \( X \). It avoids \( x_1 \) and \( x_3 \) for the same reasons \( w \) does. Now, if \( v \) is any \( m \)-factor of \( w_1 \) beginning (and ending) with \( b \), then \( v(0) \cdots v(m - 2) \sim b^{i-2} aba^{m-i-1} \).

This implies that there are \( m - i \) occurrences of \( a \) in \( v \), not situated in a contiguous block. It is now clear that \( v \not\sim x_2 \), as \( x_2 \) has only \( m - i \) \( \circ \)'s to align with the \( a \)'s, however, all appearing in a single contiguous block. Thus \( w_1 \) avoids \( x_2 \).

**Corollary 6.** Let \( X \) be an \( m \)-uniform set of three partial words of the form described in Remark 1. If \( h(x_1) + h(x_2) < m - 2 \), then \( X \) is avoidable.

**Proof.** Insert holes into \( x_1, x_2 \) so that \( 1 \leq h(x_1) \leq h(x_2), h(x_1) + h(x_2) = m - 2 \). The new set, \( X' \), is still in standard form, and is not a \( C \)-set since \( h(x_1) \geq 1 \). Hence it is avoidable by Proposition 9, and thus so is \( X \succ X' \).

Before we apply Proposition 1 to prove Theorem 1, it remains to show that the unavoidable \( C \)-sets are maximal.

**Proposition 10.** If \( m \) is even (resp., odd), then the unavoidable \( C \)-sets described in Proposition 2 (resp., Proposition 3) are maximal.

**Proof.** Let \( m \) be even, and let \( X = C_m(i) \) be an unavoidable \( C \)-set. We cannot strengthen \( x_2 \), for the resulting set would be avoidable by Corollary 6. If we strengthen \( x_3 \) with a \( b \), then the resulting set is avoidable by Proposition 9 (as it is no longer a \( C \)-set). Finally, suppose we strengthen \( x_3 \) with an \( a \) in position \( j \). Let \( i' = \min(i, j) \) and \( j' = \max(i, j) \). Then \( C_m(i', j') \) is avoidable by Proposition 3, since either \( j' \neq 2i' \), or \( j' = 2i' + m - 1 \) (since \( m - 1 \) is odd). Hence \( X \) is maximal. Now let \( m \) be odd, and let \( Y = C_m(\Lambda) \) an unavoidable \( C \)-set where \( |\Lambda| = 2 \). Again, we cannot strengthen \( x_2 \) at all, nor can we strengthen \( x_3 \) with a \( b \). Now suppose we strengthen \( x_3 \) with an
Then the resulting set is of the form $C_m(i, j, l)$, which is avoidable by Corollary 1. Hence $Y$ is maximal. 

We now complete the proof of Theorem 1.

**Proof of Theorem 1.** Let $m$ be odd (resp., even), and let $X$ be an $m$-uniform unavoidable set of three partial words, with $2m-6$ (resp., $2m-5$) total holes. Now, Corollaries 5 and 6 (resp., along with Proposition 8) together tell us that $h(x_1) + h(x_2) = m - 2$, and moreover Proposition 9 gives that $X$ is necessarily a $C$-set. But we know that unavoidable $C$-sets with $2m-6$ (resp., $2m-5$) holes are maximal, by Proposition 10, and hence $X$ is. Therefore, $H_{m,n} \geq 2m - 6$ (resp., $H_{m,n} \geq 2m - 5$) by application of Proposition 1. On the other hand, $C_m(1, 2)$ (resp., $C_m(1)$) is always unavoidable, and so we can in fact achieve $2m-6$ (resp., $2m-5$) holes in an unavoidable set. This yields the reverse inequality, that is, $H_{m,n} \leq 2m - 6$ (resp., $H_{m,n} \leq 2m - 5$).

8 Conclusion

In this paper, we have answered affirmatively a conjecture left open by Blanchet-Sadri et al. (Conjecture 2 of Reference [3]). We have computed the minimum number of holes in any unavoidable $m$-uniform set of size three over a binary alphabet (summed over all partial words in the set). We have also constructed all sets that achieve this minimum, and have shown that any unavoidable set with the stated number of holes is maximal. However, the characterization of the unavoidable sets of partial words of size three over a binary alphabet remains an open problem, even when we restrict our attention to $m$-uniform sets.

Acknowledgements

The authors would like to acknowledge Sean Simmons from the Department of Mathematics of the Massachusetts Institute of Technology for pointing out an approach to proving our characterization of the $D_m(i, j)$ unavoidable sets. We thank him for his insightful suggestion. We also thank the referees for their very valuable comments and suggestions.

A World Wide Web server interface has been established at

www.uncg.edu/cmp/research/unavoidablesets5

for automated use of a program that checks whether a given infinite word avoids a given set of three partial words of uniform length over a binary alphabet. If the answer is yes, the program outputs a shortest avoiding word.
References


