Computing the Partial Word Avoidability Indices of Binary Patterns

By: F. Blanchet-Sadri, Andrew Lohr, Shane Scott

Blanchet-Sadri, F., Lohr, A., Scott, S. (2013). Computing the Partial Word Avoidability Indices of Binary Patterns. *Journal of Discrete Algorithms, 23,* 113-118. doi: 10.1016/j.jda.2013.06.007

Abstract:

We complete the classification of binary patterns in partial words that was started in earlier publications by proving that the partial word avoidability index of the binary pattern ABABA is two and the one of the binary pattern ABBA is three.

Keywords: Combinatorics on words | Partial words | Pattern avoidance | Binary pattern | Avoidability index

Article:

1. Introduction

A *pattern* is a sequence over an alphabet of variables. An occurrence of a pattern is obtained by replacing the variables with arbitrary non-empty words, such that two occurrences of the same variable are replaced by the same word. A pattern *p* is *unavoidable* if every infinite word has an occurrence of *p*; otherwise, *p* is *avoidable*. More precisely, *p* is *k*-*unavoidable* if every infinite word over a *k*-letter alphabet has an occurrence of *p*; otherwise, *p* is *k*-*avoidable*.
The *avoidability index* of *p* is the smallest integer *k* such that *p* is *k*-avoidable (if no such integer exists, the avoidability index is ∞).

Deciding the avoidability of a pattern can be done easily [1] and [8], but deciding whether a given pattern is *k* -avoidable has remained an open problem. An alternative is the problem of classifying all the patterns over a fixed number of variables, i.e., finding the avoidability indices of all the patterns over a fixed number of variables. This problem has been completely solved for all the binary patterns, those over two variables *A* and *B* (see Chapter 3 of [7]). They fall into three categories: the patterns *ε* , *A* , *AB* , *ABA* , and their complements, are unavoidable (or have avoidability index ∞); the patterns *AA* , *AAB* , AABA, AABB,ABAB, ABBA, AABAA, AABAB, their reverses, and complements, have avoidability index 3; all other patterns, and in particular all binary patterns of length six or more, have avoidability index 2.

Recently, Blanchet-Sadri et al. [3] and [5] determined all the “non-trivial” avoidability indices of the binary patterns in *partial words* , or sequences that may have some undefined positions, called holes and denoted by ⋄’s, that match every letter of the alphabet over which they are defined (we also say that ⋄ is *compatible* with each letter of the alphabet). For example, a⋄bca⋄b is a partial word with two holes over the alphabet{a,b,c}, and aabcabb is a *full word* created by filling in the first hole with *a* and the second one with *b*. They showed that, if no variable of the pattern is substituted by a partial word consisting of only one hole, the avoidability index of the pattern remains the same as in the full word case, and they started the classification in the non-restricted to non-trivial case.

In this paper, we complete the classification of all the binary patterns that was started by Blanchet-Sadri et al., i.e., we prove that the avoidability index of the pattern ABABA is two and the one of the pattern ABBAis three. In Section 2, we give some background on partial words and patterns (for more information, see [2] and [7]) and in Section 3, we complete the classification of the avoidability indices of binary patterns.

2. Preliminaries

Let *Σ* be an *alphabet* , a non-empty finite set of symbols. Each element a∈Σ is a *letter* . A (*full*) *word* over *Σ* is a concatenation of letters from *Σ* while a *partial word* over *Σ* is a concatenation of symbols from$\Sigma_\diamond=\Sigma\cup\{\diamond\}$, the alphabet *Σ* being augmented with the “hole” symbol ⋄ (a full word is a partial word without holes). We denote by u[i] the symbol at position *i* of a partial word *u* . The *length* of *u* , |u|, is the number of symbols in *u* . The *empty word ε* is the unique word of length zero. The set of all full words (resp., non-empty full words) over *Σ* is denoted by $\Sigma^{\square}$ (resp., Σ^+), while the set of all partial words (resp., non-empty partial words) over *Σ* is denoted by $\Sigma_\diamond^*$ (resp., $\Sigma_\diamond^+$). The set of all full (resp., partial) words over *Σ* of length *n* is denoted by Σ^n (resp., $\Sigma_\diamond^n$).

A partial word *u* is a *factor* of a partial word *v* if there exist *x* , *y* such that v=xuy (the factor *u* is *proper* if u≠ε and u≠v). We say that *u* is a *prefix* of *v* if x=ε and a *suffix* of *v* if y=ε. We denote by Pref(v)the set of all prefixes of *v* and by Suf(v) the set of all suffixes of *v* . If *u* and *v* are two partial words of equal length, then *u* is *compatible* with *v* , denoted u↑v, if u[i]=v[i] whenever u[i],v[i]∈Σ. If u,vare non-empty compatible partial words, then *uv* is called a *square* . We say that *u* is compatible withPref(v) if there exists $u' \in$Pref(v) such that $u \uparrow u'$ (a similar statement holds for Suf(v)). Moreover, a full word compatible with a factor of a partial word *v* is called a *subword* of *v* . For example, ⋄b⋄ is a factor of abb⋄b⋄⋄ba and *bbb* is a subword compatible with that factor.

Let {A,B} be the binary alphabet of pattern variables with $\Sigma \cap \{A,B\} = \emptyset$. In this paper, a pattern is a word over the alphabet $\Sigma \cup \{A,B\}$. A factor $u \in \Sigma^+$ of such pattern is called a *pattern constant* . For example,*AA* is the square pattern, aAaAa is the overlap pattern, and ABBA is one of the binary patterns. For a partial word $w \in \Sigma_\diamond^*$ and pattern $p \in (\Sigma \cup \{A,B\})$□, we say that *w meets p* or *p occurs in w* if there exists some non-erasing morphism $\varphi:(\Sigma \cup \{A,B\})$□$\rightarrow \Sigma$□, which acts as the identity over Σ , such that φ(p)is compatible with a factor of *w* . We say *w avoids p* when it does not meet *p* . For example, abab meets*AA* , acbcaba avoids aAaAa, and *ababaabc⋄a⋄cd⋄⋄aba* meets ABBA. These definitions also apply to infinite partial words *w* over Σ which are functions from N to $\Sigma_\diamond$.

A pattern *p* is called *k-avoidable* if there is a partial word over a *k* -letter alphabet with infinitely many holes which avoids *p* . We say that *p* is *avoidable* if it is *k* -avoidable for some *k* . For example, *AB* is unavoidable, *AA* is unavoidable in partial words, *AA* is 3-avoidable in full words, and *AAA* is 2-avoidable [3]. For a given pattern *p* , we define its *avoidability index* as the minimal *k* such that *p* is *k* -avoidable. If *p* is unavoidable, it is ∞. For example, the avoidability indices of *AB* , AABB, and every binary pattern of length six or greater are ∞, 3, and 2, respectively [3].

3. Completion of the classification of binary patterns

The following definitions are useful for our purposes. Let Σ_1 and Σ_2 be alphabets. For a word $w \in \Sigma_2^+$ and a morphism $\varphi : \Sigma_1^* \rightarrow \Sigma_2^*$, we say that

- *w* is *φ-injected from x* if $x \in \Sigma_1^+$ is a unique word of minimal length such that *w* is a factor occurring once in φ(x) and for all $y \in \Sigma_1^+$ if *w* is a factor of φ(y) then *x* is a factor of *y*. We say *w* is *φ-injected*if such an *x* exists.

- *w* is *φ-preinjected from a* (resp., *φ-postinjected from a*) if $a \in \Sigma_1$ is such that *w* is compatible withPref(φ(a)) (resp., Suf(φ(a))).

- *w* is *φ-side-injected from a* if $a \in \Sigma_1$ is such that the number $k_a = |\{u \in \text{Pref}(\varphi(a)) | u \uparrow w\}| + |\{u \in \text{Suf}(\varphi(a)) | u \uparrow w\}|$

is exactly one, and k_b is zero for all other letters $b \in \Sigma_1$.

Let $\Sigma=\{a,b\}$, let $t:\Sigma^{*}\to\Sigma^{*}$ be the Thue–Morse morphism defined by $t(a)=ab$ and $t(b)=ba$, and let $\chi:\Sigma^{*}\to\Sigma_{\diamond}^{*}$ be the morphism defined by $\chi(a)=a$ and $\chi(b)=baaa\diamond babbb$.

Theorem 1.

The pattern ABABA*is* 2-*avoidable by* $\chi\circ t^{\omega}(a)$.

Proof.

Let Σ , t , and χ be as defined above. Assume to the contrary that $\chi\circ t^{\omega}(a)$ meets the pattern p=ABABA. Then there is some non-erasing morphism $h:\{A,B\}^{*}\to\Sigma^{*}$ and a factor w of $\chi\circ t^{\omega}(a)$ such that $h(p)\uparrow w$. It is well known that $t^{\omega}(a)$ avoids ABABA as well as overlaps and cubes [6]. We begin by noting that every factor of length five containing a hole is χ -injected. Then for any factors y , y' of $\chi\circ t^{\omega}(a)$ of at least length 5 we have that $y\uparrow y'$ implies $y=y'$.

We may write w in the form

$$w=w_1\chi(x_1)w_2|w_3\chi(x_2)w_4|w_5\chi(x_3)w_6|w_7\chi(x_4)w_8|w_9\chi(x_5)w_{10}$$

where $w_1\chi(x_1)w_2$, $w_5\chi(x_3)w_6$, and $w_9\chi(x_5)w_{10}$ are pairwise compatible, $w_3\chi(x_2)w_4$ and $w_7\chi(x_4)w_8$ are compatible, w_1 suffixes $\chi(a_1)$ for some $a_1\in\Sigma\cup\{\varepsilon\}$, w_{10} prefixes $\chi(a_6)$ for some $a_6\in\Sigma\cup\{\varepsilon\}$, and

$$w_2w_3=\chi(a_2)\quad\text{for some } a_2\in\Sigma\cup\{\varepsilon\},$$

$$w_4w_5=\chi(a_3)\quad\text{for some } a_3\in\Sigma\cup\{\varepsilon\},$$

$$w_6w_7=\chi(a_4)\quad\text{for some } a_4\in\Sigma\cup\{\varepsilon\},$$

$$w_8w_9=\chi(a_5)\quad\text{for some } a_5\in\Sigma\cup\{\varepsilon\}.$$

Note that we have inserted “|” between variable images for ease of reading.

We also allow x_i to be empty, so long as w_{2i-1}, x_i, and w_{2i} are not all simultaneously empty for any $1\leqslant i\leqslant 5$. We also choose all x_i to be maximal so that every w_i is either empty or a proper affix. Note this means that w_i is never *a*.

We see many relations of the form $u_1\chi(y_1)u_2\uparrow u_3\chi(y_2)u_4$. We consider solutions to the form for factors of $\chi\circ t^{\omega}(a)$. Every non-empty affix of $\chi(\Sigma)$ is χ -preinjected or χ -postinjected, so every pair of compatible suffixes are equal, and every pair of compatible prefixes are equal. Suppose that the lengths of the prefixes are not equal and assume towards a contradiction, and without loss of generality, that $|u_1|>|u_3|$. It is then clear that one of them must be length two or more, so to have compatible prefixes both u_1 and u_3 must be suffixes of $\chi(b)$. Then u_3 is a suffix of u_1 which must also be compatible with a prefix of u_1. The possible values of u_1 and u_3 expressed as pairs (u_1,u_3) are

$$\{(bb,b),(bbb,b),(bbb,bb),(babbb,b),(\diamond babbb,b),(\diamond babbb,bb)\}.$$

Let v be the suffix of $\chi(b)$ formed by deleting the u_3-compatible prefix from u_1. We have $v\in\{b,bb,abbb,babbb\}$. Note that a prefix of $\chi(y_2)$ must be compatible with v. No choice of the length two prefix of y_2 forms a compatible prefix of $\chi(y_2)$ for any of $\{bb,abbb,babbb\}$. So $v=b$ and b prefixes y_2. It follows that the length three prefix of y_1 is aaa. But as the Thue–Morse word avoids cubes, this cannot occur. It follows that $|u_1|=|u_3|$, and as both are either empty or suffixes of $\chi(b)$, we see that $u_1=u_3$. We can similarly show that $u_2=u_4$.
We now have that $u_1=u_3$, $u_2=u_4$, and $\chi(y_1)\uparrow\chi(y_2)$. But note that either $y_1=y_2=a$, $y_1=y_2=\varepsilon$, or $|\chi(y_1)|=|\chi(y_2)|\geqslant 10$. So $\chi(y_1)=\chi(y_2)$. But as χ is injective this yields $y_1=y_2$. We may rewrite w with fewer variables as

$$w_1\chi(x_1)w_2|w_3\chi(x_2)w_4|w_1\chi(x_1)w_2|w_3\chi(x_2)w_4|w_1\chi(x_1)w_2.$$

Because all the affixes of $\chi(\Sigma)$ are χ-preinjected or χ-postinjected, $a_1=a_3=a_5$ and $a_2=a_4=a_6$. Then w occurs only as a factor of

$$\chi(a_1x_1a_2x_2a_1x_1a_2x_2a_1x_1a_2).$$

But this yields an instance of ABABA in $t^\omega(a)$ no matter which you choose to be empty, a contradiction. Hence no factor of $\chi\circ t^\omega(a)$ is an occurrence of ABABA. □
Next, let $\Sigma=\{a,b,c\}$ and $\theta:\Sigma^{*}\to\Sigma^{*}$ be the generalized Thue–Morse morphism given by $\theta(a)=abc,\theta(b)=ac$, and $\theta(c)=b$.

Lemma 1.

The generalized Thue–Morse word $\theta^\omega(a)$ *avoids both AA and* bAbcAb.

Proof.

Assume to the contrary that $\theta^\omega(a)$ meets the pattern bAbcAb. Then there are words $x',w\in\Sigma^+$ where $w=bx'bcx'b$ is a factor of $\theta^\omega(a)$. It is well known that the fixed point $\theta^\omega(a)$ avoids squares [6]. Observe that $\theta^\omega(a)$ is also an infinite word over the alphabet {abc,ac,b}. Because *a* only occurs as a prefix in this set and *c* only as a suffix, it follows that *b* only occurs in either the factor *abc* or *cba* . Then $x'=xa$ for some $x\in\Sigma^{\square}$ and *w* is a subword of bxabcxabc. But this contains a square, which cannot appear in $\theta^{\omega}(a)$. □

Now, ABBA is 2-unavoidable for full words, which must also be true for partial words. We can prove thatABBA is 3-avoidable by considering the morphism $\varphi:\Sigma^{*}\to\Sigma_{\diamond}^{*}$ given by $\varphi(a)=cccbc,\varphi(b)=ca\diamond bcbba$, and $\varphi(c)=baa$. The proof, based on an analysis of cases, depends on Lemma 1,Lemma 2 and Lemma 3.

Lemma 2.

Let u and v be length five or greater factors of $\varphi(x)$*, with x a full word over Σ. If u and v are compatible, then they are also equal.*

Proof.

Let *u* and *v* be length five or greater compatible factors of $\varphi(x)$ with $x\in\Sigma^{+}$. We assume to the contrary that one, say *v* , has a hole in position *i* while u[i] is a letter. Note that for any word $\varphi(x)$ there are only holes in images of *b* and will be separated by at least seven letters. Then the factors u[j..j+4] andv[j..j+4] have at most one hole for any $j\leqslant|u|-5$.
If i=0 then $v[0..4]=\diamond bcbb$. But bcbb is *φ* -injected. It follows that $u[0..4]=\diamond bcbb$.
If $i=|u|-1$ then $v[|u|-3..|u|-1]=ca\diamond$. But $ca\diamond$ is *φ* -injected, so it can only be that $u[|u|-3..|u|-1]=ca\diamond$. For any other *i* we can see that $v[i-1..i+1]=a\diamond b$, but this factor is also *φ*-injected. Then no such *i* can exist and the words are equal. □

Corollary 1.

For all $x,x'\in\Sigma^{+}$*, if* $\varphi(x)$*and* $\varphi(x')$*are compatible then x and* x'*are equal.*

Proof.

If $|\varphi(x)|<5$ then $|\varphi(x)|=3$, and $x=x'=c$. Otherwise by Lemma 2 we have that $\varphi(x)=\varphi(x')$. As *φ* is injective, it is clear that $x=x'$. □

Lemma 3.

The set of square subwords of $\varphi\circ\theta^{\omega}(a)$ *is*
{aa,bb,cc,acac,baba,cbcb}.

Proof.

Let alphabet Σ={a,b,c}, set S={aa,bb,cc,acac,baba,cbcb}, and morphisms θ and φ be as defined above. Let Σ_S^n for n∈N be the set of length n square-free words of Σ^+. Naturally, as θ^ω(a) avoids squares we know that Σ_S^n contains all its subwords of length n. One may easily check that S is the set of square subwords of $\varphi(\Sigma_S^2)$ and that there are no additional squares in $\varphi(\Sigma_S^4)$. We will see that there are no other square-compatible factors of $\varphi\circ\theta^\omega$(a). Assume to the contrary that s is such a factor of $\varphi\circ\theta^\omega$(a), i.e., $s=s_1s_2$ where $s_1\uparrow s_2$. Since s is not a factor of $\varphi(\Sigma_S^4)$, it must be of the form $s=w_1\varphi(x)w_2$ for some subword x of θ^ω(a) of length four or greater and w_1 and w_2 are respectively a suffix and prefix (possibly empty) of φ(Σ). We examine cases according to which, if any, of w_1, w_2 are empty. It is evident from the possible lengths of w_1, w_2, and φ(x) that $|w_1|<|s_1|<|w_1\varphi(x)|$. So the last letter of s_1 and the first letter of s_2 occur in the image under φ of one or two adjacent letters of x. Then we may write $s=w_1\varphi(x_1)v_1v_2\varphi(x_2)w_2$ where $\varphi(x)=\varphi(x_1)v_1v_2\varphi(x_2)$ and $w_1\varphi(x_1)v_1\uparrow v_2\varphi(x_2)w_2$. Here x_1 and x_2 are non-empty subwords of θ^ω(a). We choose maximal lengths for x_1 and x_2 so that v_1, v_2 are either the empty word or there is some $a_i\in\Sigma$ with v_1 a proper prefix and v_2 a proper suffix of $\varphi(a_i)$. The length restrictions imposed on s guarantee by Lemma 2 that $w_1\varphi(x_1)v_1=v_2\varphi(x_2)w_2$. This allows us to also write s in the convenient form

$$s=w_1\varphi(y_1)u_1w_2|w_1u_2\varphi(y_2)w_2,$$

where $u_1w_2w_1u_2=\varphi(z)$ for some $z\in\Sigma\cup\Sigma_S^2$ such that $\varphi(y_1)u_1=u_2\varphi(y_2)$. Choose y_1 and y_2 of maximum length so u_1 and u_2 are empty or a proper prefix or, respectively, a suffix of φ(Σ). We proceed by considering the cases for which, if any, of w_1 and w_2 are the empty word.

Case 1. Both w_1 and w_2 are the empty word.
The square-compatible factor has the form $s=\varphi(x_1)v_1v_2\varphi(x_2)$. From Corollary 1, it is clear that v_1 and v_2 must be non-empty or we would have $x_1=x_2$ and $s=\varphi(x_1x_1)$ which would contradict the claim of Lemma 1 that θ^ω(a) contains no squares. Then $\varphi(x_1)$ must have a prefix compatible with v_2 and $\varphi(x_2)$ must have a suffix compatible with v_1, and there is some $a_i\in\Sigma$ such that $\varphi(a_i)=v_1v_2$. For any $a_i\in\Sigma$, a factorization of $\varphi(a_i)$ into v_1v_2 implies that either v_1 or v_2 is φ-side-injected, except for v_1=ba with v_2=a (*ba* is a proper suffix of φ(b) and a proper prefix of φ(c) while *a* is a proper suffix of both φ(b) and φ(c)). However, v_2 cannot equal *a* as it is not compatible with Pref(φ(Σ)). Note that v_1 must be both a proper prefix of φ(Σ) compatible with a suffix of φ(Σ) and v_2 must also be both a proper suffix of φ(Σ) compatible with a prefix of φ(Σ), which means neither is φ-side-injected. This is a contradiction.

Case 2. Both w_1 and w_2 are non-empty words.
Consider $s=w_1\varphi(y_1)u_1w_2|w_1u_2\varphi(y_2)w_2$ where $u_1w_2w_1u_2=\varphi(z)$ for some $z\in\Sigma\cup\Sigma_S^2$ such that $\varphi(y_1)u_1=u_2\varphi(y_2)$. Clearly u_1 and u_2 are both prefixes and suffixes of φ(Σ). We easily compute

the set of words both prefixing and suffixing $\varphi(\Sigma)$ to be $W=Pref(\varphi(\Sigma))\cap Suf(\varphi(\Sigma))=\{\varepsilon,c,ba\}$.

If both u_1 and u_2 are empty then $w_1w_2=\varphi(a_i)$ for some $a_i\in\Sigma$. That would mean w_1 and w_2 are also in W, but then they clearly cannot satisfy $w_2w_1=\varphi(z)$. So at least one of u_1 or u_2 must be non-empty.

Suppose u_2 is non-empty. Recall that $u_2\in W$. Then for w_1u_2 to be a suffix of $\varphi(z)$ we must have that $w_1$is a suffix of cccb,ca⋄bcb. But as w_1 is a suffix of $\varphi(\Sigma)$ it cannot end in b. Then it must be empty, a contradiction. So u_2 is empty, thus u_1 is non-empty. Then for u_1w_2 to be a prefix of $\varphi(z)$ we must havew_2 a prefix of {ccbc,a⋄bcbba,a} since $u_1\in W$. But as w_2 is also a prefix of $\varphi(\Sigma)$ we see that$w_2\in\{c,cc\}$ and the first letter of z must be a. So w_1 can only be a three- or two-letter suffix of $\varphi(a)$depending on the choice of w_2.
Then $s=cbc\varphi(x_1)cc|cbc\varphi(x_2)c$ or $s=bc\varphi(x_1)ccc|bc\varphi(x_2)cc$. But either case forces x_2 to end in a, and that forces the last letter of $\varphi(x_1)$ to be b, which is impossible. Then either w_1 or w_2 is empty.

Case 3. One of w_1,w_2 is empty.

Suppose that $w_2=\varepsilon$. We have $s=w_1\varphi(y_1)u_1|w_1u_2\varphi(y_2)$ where $u_1w_1u_2=\varphi(z)$ for some $z\in\Sigma\cup\Sigma_S^2$ and $\varphi(y_1)u_1=u_2\varphi(y_2)$. Clearly u_1 and u_2 are both prefixes and suffixes of $\varphi(\Sigma)$ and so must lie in W. Note that w_1 is a proper suffix, so both u_1 and u_2 cannot be empty or we would have $\varphi(z)=w_1$. Suppose that u_2 is non-empty. Then as w_1u_2 must suffix $\varphi(z)$ we must have$w_1\in\{cccb,ca\diamond bcb\}$, depending on the choice of u_2. But w_1 is a proper suffix of $\varphi(\Sigma)$ and neithercccb nor ca⋄bcb is such suffix. This is a contradiction. So it can only be that u_2 is empty and u_1 is non-empty, i.e., $\varphi(y_1)u_1=\varphi(y_2)$. If u_1 is c then the last letter of y_2 can only be a. But this would force$\varphi(y_1)$ to end in b, which is not a suffix of $\varphi(\Sigma)$. So $u_1=ba$. Then the last letter of y_2 must be b. But this would force $\varphi(y_1)$ to end in b. We can conclude that w_2 is not empty. The argument is symmetric if$w_1=\varepsilon$.

We have exhausted every case and we see that the only squares are those appearing as subwords of $\varphi(\Sigma_S^4)$, which we know to be *S*. □

Theorem 2.

The pattern ABBA*is* 3-*avoidable by* $\varphi\circ\theta^{\omega}(a)$.

Proof.

Let p=ABBA and let the alphabet Σ and morphisms θ and φ be as defined above. Let S be

the set of square-compatible factors of $\varphi\circ\theta^{\omega}(a)$ which has been computed in Lemma 3. Assume to the contrary that the word $\varphi\circ\theta^{\omega}(a)$ meets *p* , i.e., there is some non-erasing morphism $h:\{A,B\}^{*}\rightarrow\Sigma^{*}$ and factor *w* of $\varphi\circ\theta^{\omega}(a)$ such that $h(p)\uparrow w$.
We proceed by examining the possible instances of p=ABBA with the knowledge that h(BB)=s for some $s\in S$. Let *R* be the minimal set with every $s\in S$ a subword of $\varphi(R)$, i.e.,

$$R=\bigcup_{s\in S}\{x\in\Sigma^{*}\mid s \text{ is a subword of } \varphi(x) \text{ but not of } \varphi(y) \text{ for any proper subword } y \text{ of } x\}.$$

For $r\in R$, we write $\varphi(r)=v_1s_1s_2v_2$ where $s_1\uparrow s_2$ and v_1,v_2 are (possibly empty) affixes of $\varphi(r)$.Table 1 lists the elements of $\varphi(R)$, the corresponding square-compatible factors, and their affixes. The final column lists the affixes which are *φ*-injected. We investigate each row of the table as a separate case, but we first make some observations.

Table 1. Elements of *φ*(*R*), the corresponding square-compatible factors, and their affixes.

	φ(*r*)	v_1	s_1s_2	v_2	*φ*-injected
1	*φ*(*c*)	*b*	*aa*	*ε*	
2	*φ*(*b*)	*c*	*a⋄*	*bcbba*	*bcbba*
3	*φ*(*b*)	*ca*	*⋄b*	*cbba*	*ca, cbba*
4	*φ*(*b*)	*ca⋄bc*	*bb*	*a*	*ca⋄bc*
5	*φ*(*a*)	*ε*	*cc*	*cbc*	
6	*φ*(*a*)	*c*	*cc*	*bc*	
7	*φ*(*ab*)	*cccb*	*cc*	*a⋄bcbba*	*cccb,a⋄bcbba*
8	*φ*(*cb*)	*ba*	*aca⋄*	*bcbba*	*bcbba*
9	*φ*(*bc*)	*ca⋄bcb*	*baba*	*a*	*ca⋄bcb*
10	*φ*(*b*)	*ca*	*⋄bcb*	*ba*	*ca*
11	*φ*(*ac*)	*cc*	*cbcb*	*aa*	

Armed with *S* , it is straightforward to check for any occurrence of ABBA in $\varphi(\Sigma_S^4)$. There are none. Then we can write $w=w_1\varphi(x_1rx_2)w_2$ where $r\in R$ and w_1, w_2 are (possibly empty) suffix and prefix of $\varphi(\Sigma)$, respectively, and x_1rx_2 is a subword of $\theta^{\omega}(a)$ such that $w_1\varphi(x_1)v_1\uparrow v_2\varphi(x_2)w_2$.
By Lemma 2 we have $w_1\varphi(x_1)v_1=v_2\varphi(x_2)w_2$. When $|r|=2$, we write $r=a_ia_j$ with a_i,a_j distinct letters of *Σ*.

When $|r|=1$ consider if v_1 is *φ* -injected. Then $w_2=v_1$ is a factor of $\varphi(r)$ and $w_1\varphi(x_1)=v_2\varphi(x_2)$, and we would have that *w* is a factor of $w_1\varphi(x_1rx_2r)$. But we see this may be written as

$$w_1\varphi(x_1rx_2r)=w_1\varphi(x_1)v_1s_1s_2w_1\varphi(x_1)v_1s_1s_2v_2.$$

This would yield a square-compatible factor of $\varphi\circ\theta^{\omega}(a)$ outside of *S* , in contradiction to Lemma 3. This precludes the necessity to check Cases 3, 4, and 10. Symmetrically, *w* cannot exist if $|r|=1$ and v_2 is *φ*-injected. This precludes the necessity to check Cases 2 and 3.

If both v_1 and v_2 are φ -injected then a contradictory square in $\theta^{\omega}(a)$ is guaranteed regardless of the length of r . For if $|r|=1$ then w is a factor of $\varphi(rx_1rx_1r)$, and if $|r|=2$ then w is a factor of$\varphi(a_jx_1a_ia_jx_1a_i)$. This precludes the necessity for Cases 3 and 7.

Case 1. We have $w=w_1\varphi(x_1)b|aa|\varphi(x_2)w_2$ and $w_1\varphi(x_1)b=\varphi(x_2)w_2$. Note that for ease of reading and clarity we inserted | to separate the square-compatible factor from the rest of w . Recall that $w_2$is a prefix of $\varphi(\Sigma)$. The final letter of w_2 must be b , so $w_2\in\{b,cccb,ca\diamond b,ca\diamond bcb,ca\diamond bcbb\}$. But none of $\{ccc,ca\diamond,ca\diamond bc,ca\diamond bcb\}$ can be a suffix of $\varphi(\Sigma)$, so $w_2=b$. Then$w_1\varphi(x_1)=\varphi(x_2)$. We see that w_1 is both a prefix and suffix of $\varphi(\Sigma)$ so $w_1\in\{\varepsilon,c,ba\}$. If $w_1=\varepsilon$ we would have $x_1=x_2$ and w would be a factor of $\varphi(x_1cx_1c)$, a contradiction. If $w_1=ba$ then the first letter of x_2 is c , and we would need $\varphi(x_1)$ to be prefixed by a . It must be that $w_1=c$ and the first letter of x_2 is either a or b . Then $\varphi(x_1)$ is prefixed by ccbc or $a\diamond bcbba$, which are not in $\mathrm{Pref}(\varphi(\Sigma))$.

Case 5. We have $r=a$ and $w=w_1\varphi(x_1)|cc|cbc\varphi(x_2)w_2$. As cbc is φ -postinjected from a we have$w_1=cbc$ and $\varphi(x_1)=\varphi(x_2)w_2$. Then w is a factor of $\varphi(ax_1ax_1)$, contradicting that $\theta^{\omega}(a)$ is square-free.

Case 6. We have $r=a$, $w=w_1\varphi(x_1)c|cc|bc\varphi(x_2)w_2$, and $w_1\varphi(x_1)c=bc\varphi(x_2)w_2$. Recall that$w_1$ is a suffix of $\varphi(\Sigma)$ prefixed by bc . Then $w_1\in\{bc,bcbba\}$. But $w_1\neq bcbba$ as bba is not compatible with any prefix of $\varphi(x_2)$. Then $w_1=bc$ is a suffix of $\varphi(a)$. We have that w is a factor of$\varphi(ax_1ax_2)w_2$. Recall that w_2 is a prefix of $\varphi(\Sigma)$ ending in c . Then $w_2\in\{c,cc,ccc,ca\diamond bc\}$. If$w_2=c$ then by Corollary 1$x_1=x_2$ and w must be a factor of $\varphi(ax_1ax_1)w_2$, which implies there is a square subword of $\theta^{\omega}(a)$ contradictory to Lemma 1. We also see w_2 cannot be ccc or $ca\diamond bc$ as neithercc nor $ca\diamond b$ can suffix $\varphi(x_1)$. So $w_2=cc$ and the last letter of x_1 must be a , write $x_1=x_1'a$. But then$w$ must be a factor of $\varphi(ax_1'aax_2a)$, which implies that $\theta^{\omega}(a)$ has a square subword.

Case 8. We have $r=cb$. Note that v_2 is φ -injected from b . So

$$w=bcbba\varphi(x_1)ba|aca\diamond|bcbba\varphi(x_2)w_2$$

and $\varphi(x_1)ba=\varphi(x_2)w_2$. Recall that w_2 is a prefix of $\varphi(\Sigma)$ suffixed by ba . The only choice is$w_2=ba$. By Corollary 1 we see that $x_1=x_2$. Then w is a factor of $\varphi(bx_1cbx_1c)$. This shows a square factor in $\theta^{\omega}(a)$ contradicting Lemma 1.

Case 9. We have $r=bc$. Note that v_1 is φ -injected from b . So
$w=w_1\varphi(x_1)ca\diamond bcb|baba|a\varphi(x_2)ca\diamond bcb$

and $w_1\varphi(x_1)=a\varphi(x_2)$. Recall that w_1 is a suffix of $\varphi(\Sigma)$ prefixed by a so$w_1\in\{a,aa,a\diamond bcbba\}$. If w_1 is aa or $a\diamond bcbba$ this leaves no choice for the first letter of x_2, as neither *a* nor ⋄bcbba prefix $\varphi(\Sigma)$. We are left with the possibility that w_1=a, implying $x_1=x_2$. We see that *a* is a suffix of either $\varphi(b)$ or $\varphi(c)$. This means that *w* is a factor of $\varphi(bx_1bcx_1b)$ or$\varphi(cx_1bcx_1b)$. However either contradicts Lemma 1, which shows $\theta^{\omega}(a)$ avoids both the patternbAbcAb and squares.

Case 11. We have r=ac, $w=w_1\varphi(x_1)cc|cbcb|aa\varphi(x_2)w_2$, and $w_1\varphi(x_1)cc=aa\varphi(x_2)w_2$. Recall that w_1 is a suffix of $\varphi(\Sigma)$ beginning with *aa* . As *aa* is φ -postinjected we have w_1=aa is a suffix of $\varphi(c)$ and $\varphi(x_1)cc=\varphi(x_2)w_2$. Recall that w_2 is a prefix of $\varphi(\Sigma)$ suffixed by *cc* . Then it must be a prefix of $\varphi(a)$ and $w_2\in\{cc,ccc\}$. By Corollary 1 if w_2=cc then $x_1=x_2$, and, as w_2 only prefixes$\varphi(a)$, we see that *w* must be a factor of $\varphi(cx_1acx_1a)$, a contradictory square in $\theta^{\omega}(a)$. So w_2=ccc. Then the last letter of x_1 must have its image suffixed by *c* so $x_1=x_1'a$. We have $w=aa\varphi(x_1'aacx_2)ccc$. But this would require the square *aa* as a subword of $\theta^{\omega}(a)$ in contradiction with Lemma 1. □

Taken together with the results of [3] and [5], the complete classification of the binary patterns is summarized in the following theorem.

Theorem 3.

For partial words, binary patterns fall into three categories:

1. *The binary patterns ε, A, AA, AAB,* AABA, AABAA, *AB, ABA, and their complements, are unavoidable* (*or have avoidability index* ∞).

2. *The binary patterns* AABAB, AABB, ABAB, ABBA, *their reverses, and complements, have avoidability index* 3.

3. *All other binary patterns, and in particular all binary patterns of length six or more, have avoidability index* 2.

References

[1] D.R. Bean, A. Ehrenfeucht, G. McNulty, Avoidable patterns in strings of symbols, Pacific Journal of Mathematics 85 (1979) 261–294.

[2] F. Blanchet-Sadri, Algorithmic Combinatorics on Partial Words, Chapman & Hall/CRC Press, Boca Raton, FL, 2008.

[3] F. Blanchet-Sadri, R. Merca ¸s, S. Simmons, E. Weissenstein, Avoidable binary patterns in partial words, Acta Informatica 48 (2011) 25–41.

[4] F. Blanchet-Sadri, A. Lohr, S. Scott, Computing the partial word avoidability indices of ternary patterns, in: S. Arumugam, B. Smyth (Eds.), IWOCA 2012, 23rd International Workshop on Combinatorial Algorithms, Tamil Nadu, India, in: Lecture Notes in Computer Science, vol. 7643, Springer-Verlag, Berlin, Heidelberg, 2012, pp. 206–218.

[5] F. Blanchet-Sadri, R. Merca ¸s, S. Simmons, E. Weissenstein, Erratum to: Avoidable binary patterns in partial words, Acta Informatica 49 (2012) 53–54.

[6] M. Lothaire, Combinatorics on Words, Cambridge University Press, Cambridge, 1997.

[7] M. Lothaire, Algebraic Combinatorics on Words, Cambridge University Press, Cambridge, 2002.

[8] A.I. Zimin, Blocking sets of terms, Mathematics of the USSR. Sbornik 47 (1984) 353–364.