Periodicity on Partial Words

By: Francine Blanchet-Sadri

F. Blanchet-Sadri, "Periodicity on Partial Words." Computers and Mathematics with Applications: An International Journal, Vol. 47, No. 1, 2004, pp 71-82.

Abstract:
A partial word of length n over a finite alphabet A is a partial map from $\{0, \ldots, n - 1\}$ into A. Elements of $\{0, \ldots, n-1\}$ without image are called holes (a *word* is just a partial word without holes). A fundamental periodicity result on words due to Fine and Wilf [1] intuitively determines how far two periodic events have to match in order to guarantee a common period. This result was extended to partial words with one hole by Berstel and Boasson [2] and to partial words with two or three holes by Blanchet-Sadri and Hegstrom [3]. In this paper, we give an extension to partial words with an arbitrary number of holes.
Keywords: Combinatorial problems, Words, Formal languages.

Article:

1. INTRODUCTION

This paper relates to a fundamental periodicity result on words due to Fine and Wilf [1]. This result was extended to partial words with one, two, or three holes [2,3], and here we give an extension for an arbitrary number of holes.

Throughout the paper, i mod p denotes the remainder when dividing i by p using ordinary integer division. We also write $i \equiv j \bmod p$ to mean that i and j have the same remainder when divided by p; in other words, that p divides $i - j$ (for instance, $12 \equiv 7 \bmod 5$ but $12 \neq 7 \bmod 5$ $(2 = 7 \bmod 5)$).

1.1. Words

Let A be a nonempty finite set, or an *alphabet*. Elements of A are called *letters* and finite sequences of letters of A are called *words* over A. The unique sequence of length 0, denoted by ϵ, is called the *empty word*. The set of all words over A of finite length (greater than or equal to 0) is denoted by A^*. It is a monoid under the associative operation of concatenation or product of words (ϵ serves as identity) and is referred to as the *free monoid* generated by A. Similarly, the set of all nonempty words over A is denoted by A^+. It is a semigroup under the operation of concatenation of words and is referred to as the *free semigroup* generated by A. A word of length n over A can be defined by a map $u : \{0, \ldots, n - 1\} \to A$ but is usually represented as $u = a_0a_1 \cdots a_{n-1}$ with $a_i \in A$. The length of u or n is denoted by $|u|$.

1.2. Partial Words

Let A be a finite alphabet. A *partial word* u of length n over A is a partial map $u : \{0, \ldots, n - 1\} \to A$. If $0 < i < n$, then i belongs to the *domain* of u (denoted by Domain(u)) in case $u(i)$ is defined and i belongs to the *set of holes* of u (denoted by Hole(u)), otherwise, (a word over A is a partial word over A with an empty set of holes).

The *companion* of u (denoted by u_o) is the map $u_o : \{0, \ldots, n - 1\} \to A \cup \{o\}$ defined by

$$u_o(i) = \begin{cases} u(i), & \text{if } i \in \text{Domain}(u), \\ o, & \text{otherwise.} \end{cases}$$

The bijectivity of the map $u \longmapsto u_o$ allows us to define for partial words concepts such as concatenation in a trivial way. The symbol *o* is viewed as a "do not know" symbol and not as a "do not care" symbol as in pattern matching [2]. The word $u_o = abobbabo$ is the companion of the partial word *u* of length 8 where Domain(*u*) = {0, 1, 3, 4, 5,6} and Hole(*u*) {2, 7}.

A *period of u* is a positive integer *p* such that $u(i) = u(j)$ whenever $i, j \in$ Domain(*u*) and $i \equiv j \bmod p$ (in such a case, we call *u p-periodic*). Similarly, *a local period of u* is a positive integer *p* such that $u(i) = u(i + p)$ whenever $i, i + p \in$ Domain(*u*) (in such a case, we call *u locally p-periodic*). Every locally *p*-periodic word is *p*-periodic but not every locally *p*-periodic partial word is *p*-periodic. For instance, the partial word with companion *aboaoaaa* is locally three-periodic but is not three-periodic.

2. PERIODICITY

In this section, we discuss periodicity results on partial words with zero, one, two, or three holes.

2.1. On Partial Words with Zero or One Hole

In this section, we restrict ourselves to partial words with zero or one hole.

THEOREM 1. *(See [1,2].) Let p and q be positive integers.*

(1) *Let u be a word. If u is p-periodic and q-periodic and* $|u| \geq p + q — \gcd(p, q)$*, then u is* gcd*(p, q)-periodic.*

(2) *Let u be a partial word such that card(*Hole*(u))* = 1*. If u is locally p-periodic and locally q-periodic and* $|u| \geq p + q$*, then u is* gcd*(p, q)-periodic.*

The bound $p + q — \gcd(p, q)$ turns out to be optimal in Theorem 1(1). For example, the word *abaaba* of length 6 is three-periodic and five-periodic but is not one-periodic. Also, the bound $p + q$ is optimal in Theorem 1(2) as can be seen with *abaabao* of length 7 which is locally three-periodic and locally five-periodic but not one-periodic.

2.2. On Partial Words with Two or Three Holes

In [3], it was shown that the concept of (2,*p, q*)-special (respectively, (3,*p, q*)-special) partial word is crucial for extending Theorem 1 to two holes (respectively, three holes).

DEFINITION 1. *(See [3].) Let p and q be positive integers satisfying* $p < q$*. A partial word u is called*

(1) *(2, p, q)-special if at least one of the following holds.*

(a) $q = 2p$ *and there exists* $p \leq i < |u| — 4p$ *such that* $i + p, i + 2p \in$ Hole*(u).*

(b) *There exists* $0 \leq i < p$ *such that* $i + p, i + q \in$ Hole*(u).*

(c) *There exists* $|u| — p \leq i < |u|$ *such that* $i — p, i — q \in$ Hole*(u).*

(2) *(3,p, q)-special if it is (2,p, q)-special or if at least one of the following holds.*

(a) $q = 3p$ *and there exists* $p \leq i < |u| - 5p$ *such that* $i + p, i + 2p, i + 3p \in$ Hole*(u) or there exists* $p \leq i < |u| - 7p$ *such that* $i + p, i + 3p, i + 5p \in$ Hole*(u).*

(b) *There exists* $0 \leq i < p$ *such that* $i + q, i + 2p, i +p+ q \in$ Hole*(u).*

(c) *There exists* $|u| - p \leq i < |u|$ *such that* $i - q, i - 2p, i-p- q \in$ Hole*(u).*

(d) *There exists* $p \leq i < q$ *such that* $i - p, i + p, i + q \in$ Hole*(u).*

(e) *There exists* $|u| - q \leq i < |u| - p$ *such that* $i - p, i + p, i - q \in$ Hole*(u).*

(f) $2q = 3p$ *and there exists* $p \leq i < |u| - 5p$ *such that* $i + q, i + 2p, i +p + q \in$ Hole*(u).*

If *p* and *q* are positive integers satisfying $p < q$ and $\gcd(p, q) = 1$, then the infinite sequence $(ab^{p-1}ob^{q-p-1}ob^n)_{n>0}$ consists of binary (2,*p, q*)-special partial words with two holes that are locally *p*-periodic and locally *q*-periodic but not one-periodic. Similarly, the infinite sequence $(oab^{p-1}obq\text{-}p\text{-}1ob^n)_{n>0}$ consists of binary (3,*p, q*)-special partial words with three holes that are locally *p*-periodic and locally *q*-periodic but not one-periodic.

THEOREM 2. *(See [3].) Let p and q be positive integers satisfying* $p < q$.

(1) *Let u be a partial word such that* card(Hole(u)) $= 2$ *and assume that u is not* (2,*p*,*q*)*special. If u is locally p-periodic and locally q-periodic and* $|u| \geq 2(p + q) - \gcd(p, q)$, *then u is* gcd(*p*, *q*)*-periodic.*

(2) *Let u be a partial word such that* card(Hole(u)) $= 3$ *and assume that u is not* (3, *p*, *q*)*-special. If u is locally p-periodic and locally q-periodic and* $|u| \geq 2(p + q)$, *then u is* gcd(*p*, *q*)*-periodic.*

The bound $2(p + q) - \gcd(p, q)$ turns out to be optimal in Theorem 2(1). For instance, the partial word with companion *abaabaooabaaba* of length 14 is locally three-periodic and locally five-periodic but is not one-periodic. A similar result holds for the bound $2(p + q)$ in Theorem 2(2) by considering *abaabaooabaabao*.

3. SPECIAL PARTIAL WORDS

In this section, we give an extension of the notions of (2,*p*, *q*)- and (3,*p*, *q*)-special partial words. We first discuss the case where $p = 1$ and then the case where $p > 1$.

3.1. p = 1

Throughout this section, we fix $p = 1$. Let q be an integer satisfying $q > 1$. Let u be a partial word of length $\underline{n}$ that is locally p-periodic and locally q-periodic. The companion of u, $u_o = u_o(0)u_o(1) \ldots u_o(n - 1)$, can be represented as a two-dimensional structure in the following fashion:

$$\begin{array}{cccc} u_o(0) & u_o(q) & u_o(2q) & \cdots \\ u_o(1) & u_o(1+q) & u_o(1+2q) & \cdots \\ \vdots & \vdots & \vdots & \\ u_o(q-1) & u_o(2q-1) & u_o(3q-1) & \cdots \end{array}.$$

If we wrap the array around and sew the last row to the first row so that $u_o(q - 1)$ is sewn to $u_o(q)$, $u_o(2q - 1)$ is sewn to $u_o(2q)$, and so on, then we get a cylinder for u_o.

We say that $i - p$ (respectively, $i + p$) is immediately *above* (respectively, *below*) i whenever $p \leq i < n$ (respectively, $0 \leq i < n - p$). Similarly, we say that $i - q$ (respectively, $i + q$) is immediately *left* (respectively, *right*) of i whenever $q \leq i < n$ (respectively, $0 \leq i < n - q$). The fact that u is locally p-periodic implies that if i, $i + q \in$ Domain(u), then $u(i) = u(i+p)$. Similarly, the fact that u is locally q-periodic implies that if i, $i + q \in$ Domain(u), then $u(i) = u(i + q)$.

The following define three types of isolation that will be acceptable in our definition of special partial word. In Type 1, we have a continuous sequence of holes isolating a subset of defined positions (this type of isolation occurs at the beginning of the partial word). In Type 2, a continuous sequence of holes completely surrounds a subset of defined positions. Finally, in Type 3, a continuous sequence of holes isolates a subset of defined positions (this type of isolation occurs at the end of the partial word).

DEFINITION 2. *Let S be a nonempty proper subset of* Domain(u). *We say that* Hole(u) *one-isolates S (or that S is one-isolated by* Hole(u)) *if the following hold.*

(1) *Left: if* $i \in S$ *and* $i \geq q$, *then* $i—q \in S$ *or* $i—q \in$ Hole(u).

(2) *Right: if* $i \in S$, *then* $i+q \in S$ *or* $i + q \in$ Hole(u).

(3) *Above: if* $i \in S$ *and* $i \geq p$, *then* $i—p \in S$ *or* $i— p \in$ Hole(u).

(4) *Below: if* $i \in S$, *then* $i+p \in S$ *or* $i+ p \in$ Hole(u).

DEFINITION 3. *Let S be a nonempty proper subset of* Domain(u). *We say that* Hole(u) *two-isolates S (or that S is two-isolated by* Hole(u)) *if the following hold.*

(1) *Left: if* $i \in S$, *then* $i — q \in S$ *or* $i — q \in$ Hole(u).

(2) *Right: if* $i \in S$, *then* $i + q \in S$ *or* $i + q \in$ Hole(u).

(3) *Above: if* $i \in S$, *then* $i — p \in S$ *or* $i — p \in$ Hole(u).

(4) *Below: if* $i \in S$, *then* $i + p \in S$ *or* $i + p \in$ Hole(u).

DEFINITION 4. *Let S be a nonempty proper subset of* Domain*(u). We say that* Hole*(u) three-isolates S (or that S is three-isolated by* Hole*(u)) if the following hold.*

(1) *Left: if $i \in S$, then $i - q \in S$ or $i - q \in$* Hole*(u).*
(2) *Right: if $i \in S$ and $i < n - q$, then $i + q \in S$ or $i + q \in$* Hole*(u).*
(3) *Above: if $i \in S$, then $i - p \in S$ or $i - p \in$* Hole*(u).*
(4) *Below: if $i \in S$ and $i < n - p$, then $i + p \in S$ or $i + p \in$* Hole*(u).*

EXAMPLE 1. As a first example, consider the partial word u_1 with companion $(u_1)_o$ represented as the two-dimensional structure of Figure 1. Here, u_1 is locally one-periodic and locally five-periodic.

	0	5	10	15	20	25	30	35	40	45	50	55	60
0	c	o	a	a	a	o	d	o	e	o	f	f	o
1	c	o	o	a	o	h	o	e	o	f	f	o	i
2	o	o	b	o	o	o	e	e	e	o	f	f	o
3	a	o	b	b	o	e	e	e	o	f	f	f	f
4	a	a	o	o	g	o	e	e	e	o	f	f	

Figure 1.

The set of positions with letter a is one-isolated by Hole(u_1); the set of positions with letter b is two-isolated by Hole(u_1); the set of positions with letter c is one-isolated by Hole(u_1); the set of positions with letter d is two-isolated by Hole(u_1); the set of positions with letter e is two-isolated by Hole(u_1); the set of positions with letter f is three-isolated by Hole(u_1); the set of positions with letter g is two-isolated by Hole(u_1); the set of positions with letter h is two-isolated by Hole(u_1); the set of positions with letter i is three-isolated by Hole(u_1).

EXAMPLE 2. As a second example, consider the locally one-periodic and locally five-periodic partial word u_2 with companion $(u_2)o$ represented as the two-dimensional structure of Figure 2. We can see that Domain(u_2) does not contain a nonempty subset of isolated positions.

DEFINITION 5. *Let q be an integer satisfying $q > 1$. For $1 \leq i \leq 3$, the partial word u is called* (card*(*Hole*(u)), 1, q)-special of type i if* Hole*(u) i-isolates a nonempty proper subset of* Domain*(u). The partial word u is called* (card*(*Hole*(u)), 1, q)-special if u is* (card*(*Hole*(u)), 1, q)-special of type i for some $i \in \{1, 2,3\}$.*

	0	5	10	15	20	25	30	35	40	45	50	55	60
0	a	a	a	a	a	o	a	a	a	o	a	a	o
1	a	o	o	a	o	a	o	a	o	a	a	o	a
2	o	o	a	a	o	a	a	a	a	a	a	a	a
3	a	o	a	a	a	a	a	a	o	a	a	a	a
4	a	a	o	o	a	o	a	a	a	o	a	a	

Figure 2.

It is a simple matter to check that the above definition extends the notion of (2,1, q)-special and the notion of (3, 1, q)-special (as given in Definition 1). Definition 1(1) corresponds to arrays like

(a) $q = 2$

$$\begin{array}{cccccccc} u_o(0) & u_o(2) & \cdots & u_o(2m) & o & u_o(4+2m) & \cdots \\ u_o(1) & u_o(3) & \cdots & u_o(1+2m) & o & u_o(5+2m) & \cdots \end{array}$$

or

$$\begin{array}{cccccccc} u_o(0) & u_o(2) & \cdots & u_o(2m) & u_o(2+2m) & o & u_o(6+2m) & \cdots \\ u_o(1) & u_o(3) & \cdots & u_o(1+2m) & o & u_o(5+2m) & u_o(7+2m) & \cdots \end{array}$$

(b)

$$\begin{array}{cccc} u_o(0) & o & u_o(2q) & \cdots \\ o & u_o(1+q) & u_o(1+2q) & \cdots \\ u_o(2) & u_o(2+q) & u_o(2+2q) & \cdots \\ \vdots & \vdots & \vdots & \\ u_o(q-1) & u_o(2q-1) & u_o(3q-1) & \cdots \end{array}$$

and the symmetrical of (b) for Definition 1(1)(c). Similarly, Definition 1(2) corresponds to arrays like

(a) $q = 3$

$$\begin{array}{cccccc} u_o(0) & u_o(3) & \cdots & u_o(3m) & o & u_o(6+3m) & \cdots \\ u_o(1) & u_o(4) & \cdots & u_o(1+3m) & o & u_o(7+3m) & \cdots \\ u_o(2) & u_o(5) & \cdots & u_o(2+3m) & o & u_o(8+3m) & \cdots \end{array}$$

or

$$\begin{array}{cccccccc} u_o(0) & u_o(3) & \cdots & u_o(3m) & u_o(3+3m) & o & u_o(9+3m) & \cdots \\ u_o(1) & u_o(4) & \cdots & u_o(1+3m) & u_o(4+3m) & o & u_o(10+3m) & \cdots \\ u_o(2) & u_o(5) & \cdots & u_o(2+3m) & o & u_o(8+3m) & u_o(11+3m) & \cdots \end{array}$$

or

$$\begin{array}{cccccccc} u_o(0) & u_o(3) & \cdots & u_o(3m) & u_o(3+3m) & o & u_o(9+3m) & \cdots \\ u_o(1) & u_o(4) & \cdots & u_o(1+3m) & o & u_o(7+3m) & u_o(10+3m) & \cdots \\ u_o(2) & u_o(5) & \cdots & u_o(2+3m) & o & u_o(8+3m) & u_o(11+3m) & \cdots \end{array}$$

or

$$\begin{array}{cccccccc} u_o(0) & u_o(3) & \cdots & u_o(3m) & o & u_o(6+3m) & u_o(9+3m) & \cdots \\ u_o(1) & u_o(4) & \cdots & u_o(1+3m) & u_o(4+3m) & o & u_o(10+3m) & \cdots \\ u_o(2) & u_o(5) & \cdots & u_o(2+3m) & o & u_o(8+3m) & u_o(11+3m) & \cdots \end{array}$$

or

$$\begin{array}{cccccccc} u_o(0) & u_o(3) & \cdots & u_o(3m) & u_o(3+3m) & o & u_o(9+3m) & \cdots \\ u_o(1) & u_o(4) & \cdots & u_o(1+3m) & o & u_o(7+3m) & u_o(10+3m) & \cdots \\ u_o(2) & u_o(5) & \cdots & u_o(2+3m) & u_o(5+3m) & o & u_o(11+3m) & \cdots \end{array}$$

or

$$\begin{array}{ccccccccc} u_o(0) & u_o(3) & \cdots & u_o(3m) & u_o(3+3m) & u_o(6+3m) & o & u_o(12+3m) & \cdots \\ u_o(1) & u_o(4) & \cdots & u_o(1+3m) & u_o(4+3m) & o & u_o(10+3m) & u_o(13+3m) & \cdots \\ u_o(2) & u_o(5) & \cdots & u_o(2+3m) & o & u_o(8+3m) & u_o(11+3m) & u_o(14+3m) & \cdots \end{array}$$

(and similarly for Definition 1(2)(f)) and

(b)

$$\begin{array}{cccc} u_o(0) & o & u_o(2q) & \cdots \\ u_o(1) & o & u_o(1+2q) & \cdots \\ o & u_o(2+q) & u_o(2+2q) & \cdots \\ u_o(3) & u_o(3+q) & u_o(3+2q) & \cdots \\ \vdots & \vdots & \vdots & \\ u_o(q-1) & u_o(2q-1) & u_o(3q-1) & \cdots \end{array}$$

(d)

$$\begin{array}{cccc} o & u_o(q) & u_o(2q) & \cdots \\ u_o(1) & o & u_o(1+2q) & \cdots \\ o & u_o(2+q) & u_o(2+2q) & \cdots \\ u_o(3) & u_o(3+q) & u_o(3+2q) & \cdots \\ \vdots & \vdots & \vdots & \\ u_o(q-1) & u_o(2q-1) & u_o(3q-1) & \cdots \end{array}$$

$$\vdots$$

$$\begin{array}{cccc} u_o(0) & u_o(q) & u_o(2q) & \cdots \\ u_o(1) & u_o(1+q) & u_o(1+2q) & \cdots \\ \vdots & \vdots & \vdots & \\ u_o(q-4) & u_o(2q-4) & u_o(3q-4) & \cdots \\ o & u_o(2q-3) & u_o(3q-3) & \cdots \\ u_o(q-2) & o & u_o(3q-2) & \cdots \\ o & u_o(2q-1) & u_o(3q-1) & \cdots \end{array}$$

$$\begin{array}{cccc} u_o(0) & o & u_o(2q) & \cdots \\ u_o(1) & u_o(1+q) & u_o(1+2q) & \cdots \\ \vdots & \vdots & \vdots & \\ u_o(q-3) & u_o(2q-3) & u_o(3q-3) & \cdots \\ o & u_o(2q-2) & u_o(3q-2) & \cdots \\ u_o(q-1) & o & u_o(3q-1) & \cdots \end{array}$$

and the symmetrical of (b) for Definition 1(2)(c) as well as the symmetrical of (d) for Definition 1(2)(e).

We can also check that the partial word u_1 depicted in Figure 1 is (25, 1, 5)-special, but the partial word u_2 depicted in Figure 2 is not (18, 1, 5)-special.

3.2. p > 1

Throughout this section, we fix $p > 1$. Let q be an integer satisfying $p < q$. Let u be a partial word of length n that is locally p-periodic and locally q-periodic. We illustrate with examples how the positions of the companion of u can be represented as a two-dimensional structure.

In a case where $\gcd(p, q) = 1$ (like $p = 2$ and $q = 5$), we get one array

$$\begin{array}{cccccc}
 & u_o(0) & u_o(5) & u_o(10) & u_o(15) & \cdots \\
 & u_o(2) & u_o(7) & u_o(12) & u_o(17) & \cdots \\
 & u_o(4) & u_o(9) & u_o(14) & u_o(19) & \cdots \\
u_o(1) & u_o(6) & u_o(11) & u_o(16) & u_o(21) & \cdots \\
u_o(3) & u_o(8) & u_o(13) & u_o(18) & u_o(23) & \cdots .
\end{array}$$

If we wrap the array around and sew the last row to the first row so that $u_o(3)$ is sewn to $u_o(5)$, $u_o(8)$ is sewn to $u_o(10)$, and so on, then we get a cylinder for the positions of u_o.

In a case where $\gcd(p, q) = 2$ (like $p = 6$ and $q = 8$), we get two arrays

$$\begin{array}{ccccccc}
 & & u_o(0) & u_o(8) & u_o(16) & u_o(24) & \cdots \\
 & & u_o(6) & u_o(14) & u_o(22) & u_o(30) & \cdots \\
 & u_o(4) & u_o(12) & u_o(20) & u_o(28) & u_o(36) & \cdots \\
u_o(2) & u_o(10) & u_o(18) & u_o(26) & u_o(34) & u_o(42) & \cdots
\end{array}$$

and

$$\begin{array}{ccccccc}
 & & u_o(1) & u_o(9) & u_o(17) & u_o(25) & \cdots \\
 & & u_o(7) & u_o(15) & u_o(23) & u_o(31) & \cdots \\
 & u_o(5) & u_o(13) & u_o(21) & u_o(29) & u_o(37) & \cdots \\
u_o(3) & u_o(11) & u_o(19) & u_o(27) & u_o(35) & u_o(43) & \cdots .
\end{array}$$

If we wrap the first array around and sew the last row to the first row so that $u_o(2)$ is sewn to $u_o(8)$, $u_o(10)$ is sewn to $u_o(16)$, and so on, then we get a cylinder for some of the positions of u_o. The other positions are in the second array where we wrap around and sew the last row to the first row so that $u_o(3)$ is sewn to $u_o(9)$, $u_o(11)$ is sewn to $u_o(17)$, and so on.

In general, if $\gcd(p, q) = d$, we get d arrays. In this case, we say that $i - p$ (respectively, $i + p$) is immediately *above* (respectively, *below*) i (within one of the d arrays) whenever $p \leq i < n$ (respectively, $0 \leq i < n - p$). Similarly, we say that $i - q$ (respectively, $i + q$) is immediately *left* (respectively, *right*) of i (within one of the d arrays) whenever $q \leq i < n$ (respectively, $0 \leq i < n - q$). As before, the fact that u is locally p-periodic implies that if $i, i + p \in \text{Domain}(u)$, then $u(i) = u(i + p)$. Similarly, the fact that u is locally q-periodic implies that if $i, i + q \in \text{Domain}(u)$, then $u(i) = u(i + q)$.

In what follows, we define $N_j = \{i \mid i \geq 0 \text{ and } i \equiv j \bmod \gcd(p, q)\}$ for $0 \leq j < \gcd(p, q)$.

DEFINITION 6. *Let p and q be positive integers satisfying p < q. For $1 \leq i \leq 3$, the partial word u is called* (card(Hole(*u*)),*p, q*)*-special of type i if there exists $0 \leq j <$* gcd(*p, q*) *such that* Hole(*u*) *i-isolates a nonempty proper subset of* Domain(*u*) $\cap N_j$. *The partial word u is called* (card(Hole(*u*)),*p, q*)*-special if u is* (*card*(Hole(*u*)), *p, q*)*-special of type i for some $i \in \{1, 2, 3\}$.*

EXAMPLE 3. As a first example, the partial word u_3 of Figure 3 is (5, 2, 5)-special ($p = 2$ and $q = 5$). The set of positions {0,2,4,9} is one-isolated by Hole(u_3).

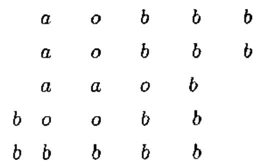
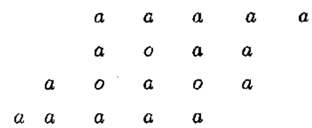


Figure 3.

EXAMPLE 4. As a second example, the partial word u_4 of Figure 4 is not (6, 6, 8)-special.



and

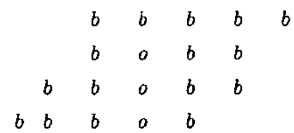


Figure 4.

4. GRAPHS ASSOCIATED WITH PARTIAL WORDS

Let p and q be positive integers satisfying $p < q$. In this section, we associate to a partial word u that is locally p-periodic and locally q-periodic an undirected graph $G_{(p,q)}(u)$. Whether or not u is (card(Hole(u)), p, q)-special will be seen from $G_{(p,q)}(u)$.

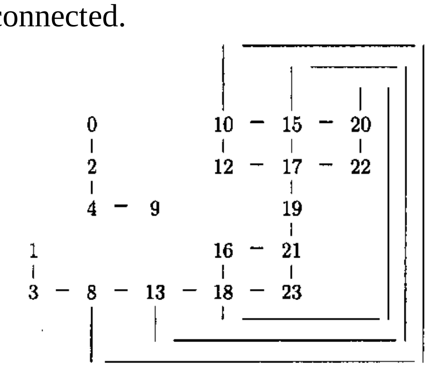
As explained in Section 3, the companion of u, $u_o = u_o(0)u_o(1) \ldots u_o(|u|-1)$, can be represented as a two-dimensional structure. Each of the gcd(p, q) arrays of u is associated with a subgraph $G = (V, E)$ of $G_{(p,q)}(u)$ as follows.

V is the subset of Domain(u) comprising the defined positions of u within the array, $E = E_1 \cup E_2 \cup E_3 \cup E_4$ where

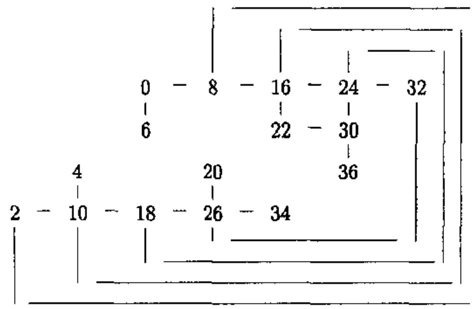
$$\begin{aligned} E_1 &= \{\{i, i-q\} \mid i, i-q \in V\}, \\ E_2 &= \{\{i, i+q\} \mid i, i+q \in V\}, \\ E_3 &= \{\{i, i-p\} \mid i, i-p \in V\}, \\ E_4 &= \{\{i, i+p\} \mid i, i+p \in V\}. \end{aligned}$$

For $0 \leq j < \gcd(p, q)$, the subgraph of $G_{(p,q)}(u)$ corresponding to Domain(u) $\cap N_j$ will be denoted by $G^j_{(p,q)}(u)$. Whenever gcd(p, q) = 1, $G^0_{(p,q)}(u)$ is just $G_{(p,q)}(u)$.

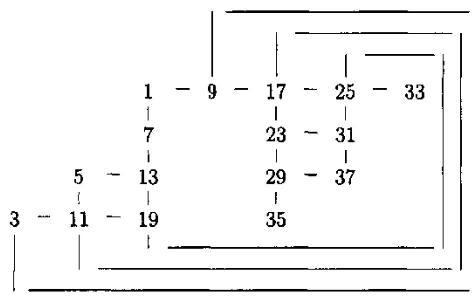
EXAMPLE 5. As a first example, the graph of the partial word u_3 of Figure 3, $G_{(2,5)}(u_3)$, is shown in the following figure and is seen to be disconnected.



EXAMPLE 6. As a second example, consider the partial word u_4 of Figure 4. The subgraphs of $G_{(6,8)}(u_4)$ corresponding to the two arrays of u_4, $G^0_{(6,8)}(u_4)$ and $G^1_{(6,8)}(u_4)$, are shown below and are seen to be connected.



and



We now define the critical lengths. We consider an even number of holes $2N$ and an odd number of holes $2N + 1$.

DEFINITION 7. Let p and q be positive integers satisfying $p < q$. The *critical lengths* for p and q are defined as follows:

(1) $\ell_{(2N,p,q)} = (N + 1)(p + q) - \gcd(p, q)$ for $N \geq 0$, and
(2) $\ell_{(2N+1,p,q)} = (N + 1)(p + q)$ for $N \geq 0$.

LEMMA 1. *Let p and q be positive integers satisfying $p < q$, and let H be a positive integer. Let u be a partial word such that* card(Hole(u)) $= H$ *and assume that* $|u| > \ell_{(H,p,q)}$. *Then u is not (H, p, q)-special if and only if $G^j_{(p,q)}(u)$ is connected for all $0 \leq j < \gcd(p, q)$.*

PROOF. We first show that if u is (H, p, q)-special, then there exists $0 \leq j < \gcd(p, q)$ such that $G^j_{(p,q)}(u)$ is not connected. Three cases arise.

CASE 1. u is (H, p, q)-special of Type 1.
There exists $0 \leq j < \gcd(p, q)$ such that Hole(u) one-isolates a nonempty proper subset S of Domain(u) $\cap N_j$. The subgraph of $G^j_{(p,q)}(u)$ with vertex set S constitutes a union of components (one component or more). There are, therefore, at least two components in $G^j_{(p,q)}(u)$ since S is proper.

CASE 2. u is (H, p, q)-special of Type 2.
This case is similar to Case 1.

CASE 3. u is (H, p, q)-special of Type 3.
This case is similar to Case 1.

We now show that if there exists $0 \leq j < \gcd(p, q)$ such that $G^j_{(p,q)}(u)$ is not connected, then u is (H, p, q)-special (or Hole(u) isolates a nonempty proper subset of Domain(u) $\cap N_j$). Consider such a j. Put $p = p' \gcd(p, q)$ and $q = q' \gcd(p, q)$. As before, the partial word u_j is defined by

$$(u_j)_o = u_o(j)u_o(j + \gcd(p,q))u_o(j + 2\gcd(p,q)) \cdots .$$

If $H = 2N$ for some N, u_j, is of length at least $(N + 1)(p' + q') — 1$; and if $H = 2N + 1$ for some N, u_j is of length at least $(N + 1)(p' + q')$. In order to simplify the notation, let us denote $G_{(p',q')}(u_j)$ by G^j. Our assumption implies that G^j is not connected.

(1) Let G_o^j be the graph constructed for the companion word $(u_j)_o$, so there are no holes. Then G^j is a subgraph of G_o^j obtained by removing the "hole" vertices.

(2) Consider a set of consecutive indices in the domain of $(u_j)_o$, say $i, i + \gcd(p, q), \ldots, i + n \gcd(p, q)$. Call such a set a "domain interval", of length $n + 1$.

(3) Every domain interval of length $p' + q'$ is the set of vertices of a cycle in G_o^j; that is, there is a closed path in G_o^j which goes through exactly this set of vertices. The point is that a cycle cannot be disconnected by just one point.

(4) Suppose C and C' are components of G^j with vertex sets S and S', and suppose neither S nor S' is isolated. Then each domain interval of length $p' + q'$ must contain a point v from S and a point v' from S'.

(5) There must be two holes in each domain interval of length $p' + q'$, since otherwise the points v and v' from Item 4 would be connected by a path in the cycle formed by the domain interval.

(6) If the number of holes is $2N + 1$ and the length of $(u_j)o$ is at least $(N + 1)(p' + q')$, then Item 5 is impossible, since $(u_j)o$ would have $N + 1$ pairwise disjoint domain intervals of length $p' + q'$ and Item 5 would then require $2(N + 1)$ holes. Similarly, if the number of holes is $2N$ and the length of $(u_j)o$ is at least $(N +1)(p' + q') —1$, then Item 5 is impossible since $(u_j)o$ would have N pairwise disjoint intervals of length $p' + q'$ and one remaining of length $p' + q' — 1$, and so Item 5 would require $2N + 1$ holes.

Note that this proves the lemma in case the number of holes is positive, and in fact Item 3 is essentially the proof in the case of exactly one hole. The case of zero holes follows from the fact that every domain interval of length $p' + q' — 1$ is the set of vertices of a path in G_o^j.

5. AN ARBITRARY NUMBER OF HOLES

In this section, we give our main result which extends Theorems 1 and 2 to an arbitrary number of holes.

LEMMA 2. *Let p and q be positive integers satisfying $p < q$ and* $\gcd(p, q) = 1$. *Let u be a partial word that is locally p-periodic and locally q-periodic. If $G_{(p,q)}(u)$ is connected, then u is one-periodic.*

PROOF. Let i be a fixed position in Domain(u). If $i' \in$ Domain(u) and $i' \neq i$, then there is a path in $G_{(p,q)}(u)$ between i' and i. Let $i', i_1, i_2, \ldots, i_n, i$ be such a path. We get $u(i') = u(i_1) = u(i_2) = \ldots = u(i_n) = u(i)$.

THEOREM 3. *Let p and q be positive integers satisfying $p < q$. Let u be a partial word that is locally p-periodic and locally q-periodic. If $G^i_{(p,q)}(u)$ is connected for all $0 \le i <$ $\gcd(p, q)$, then u is* $\gcd(p, q)$*-periodic.*

PROOF. The case where $\gcd(p, q) = 1$ follows by Lemma 2. So consider the case where $\gcd(p, q) > 1$. Define for each $0 \le i < \gcd(p, q)$ the partial word u_i by

$$(u_i)o = u_o(i)u_o(i + \gcd(p, q))u_o(i + 2\gcd(p, q)) \ldots .$$

Put $p = p' \gcd(p, q)$ and $q = q' \gcd(p, q)$. Each u_i is locally p'-periodic and locally q'-periodic. If $G^i_{(p,q)}(u)$ is connected for all i, then $G_{(p',q')}(u_i)$ is connected for all i. Consequently, each u_i is one-periodic by Lemma 2, and u is $\gcd(p, q)$-periodic.

THEOREM 4. *Let p and q be positive integers satisfying $p < q$, and let H be a positive integer. Let u be a* *partial word such that* card(Hole$(u)) = H$ *and assume that* u *is not* (H, p, q)*-special. If* u *is locally* p*-periodic and locally* q*-periodic and* $|u| \geq \ell_{(H,p,q)}$*, then* u *is* gcd(p, q)*-periodic.*

PROOF. If u is not (H, p, q)-special and $|u| \geq \ell_{(H,p,q)}$, then $G^{i}_{(p,q)}(u)$ is connected for all $0 \leq i < \gcd(p, q)$ by Lemma 1. Then u is gcd(p, q)-periodic by Theorem 3.

The bound $\ell_{(2N,p,q)}$ turns out to be optimal for an even number of holes $2N$, and the bound $\ell_{(2N+1,p,q)}$ optimal for an odd number of holes $2N+1$. The following builds a sequence of partial words showing this optimality.

DEFINITION 8. *Let* p *and* q *be positive integers satisfying* $1 < p < q$ *and* gcd$(p, q) = 1$*. Let* N *be a positive integer.*

(1) *The partial word* $u_{(2N,p,q)}$ *over* $\{a, b\}$ *of length* $\ell_{(2N,p,q)} - 1$ *is defined by*

(a) Hole$(u_{(2N,p,q)}) = \{p+q-2, p+q-1, 2(p+q)-2, 2(p+q)-1, \ldots, N(p+q)-2, N(p+q)-1\}$.

(b) *The component of the graph* $G_{(p,q)}(u)$ *containing* $p-2$ *is colored with letter* a.

(c) *The component of the graph* $G_{(p,q)}(u)$ *containing* $p-1$ *is colored with letter* b.

(2) *The partial word* $u_{(2N+1,p,q)}$ *over* $\{a, b\}$ *of length* $\ell_{(2N+1,p,q)} - 1$ *is defined by* $(u_{(2N+1,p,q)})_o = u_{(2N,p,q)}o$ *so that* Hole$(u_{(2N+1,p,q)}) =$ Hole$(u_{(2N,p,q)}) \cup \{(N+1)(p+q)-2\}$.

The partial word $u_{(2N,p,q)}$ can be thought of as two bands of holes $\text{Band}_1 = \{p+q-1, 2(p+q)-1, \ldots, N(p+q)-1\}$ and $\text{Band}_2 = \{p+q-2, 2(p+q)-2, \ldots, N(p+q)-2\}$ where between the bands the letter is a and outside the bands it is b or vice versa (a similar statement holds for $u_{(2N+1,p,q)}$).

EXAMPLE 7. For example, the partial word $u_{(4,2,5)}$ of length 19 has companion represented as the two-dimensional structure:

	a	o	b	b
	a	a	o	b
	a	a	a	
b	o	a	a	
b	b	o	a	

It is locally two-periodic and locally five-periodic but is not one-periodic (it is not (4, 2, 5)-special).

EXAMPLE 8. Similarly, the partial word $u_{(5,2,5)}$ of length 20 has companion represented as the two-dimensional structure:

	a	o	b	b
	a	a	o	b
	a	a	a	o
b	o	a	a	
b	b	o	a	

It is locally two-periodic and locally five-periodic but is not one-periodic (it is not (5, 2, 5)-special).

PROPOSITION 1. *Let* p *and* q *be positive integers satisfying* $1 < p < q$ *and* gcd$(p, q) = 1$. *Let* H *be a positive integer. The partial word* $u_{(H,p,q)}$ *of length* $\ell_{(H,p,q)} - 1$ *is not* (H, p, q)*-special, but is locally* p*-periodic and locally* q*-periodic. However,* $u_{(H,p,q)}$ *is not one-periodic.*

PROOF. We prove the result when $H = 2N$ for some N (the odd case $H = 2N + 1$ is similar). As stated earlier, the partial word $u_{(2N,p,q)}$ of length $(N+1)(p+q) - 2$ can be thought of as two bands of holes $\text{Band}_1 = \{p+q-1, 2(p+q)-1, \ldots, N(p+q)-1\}$ and $\text{Band}_2 = \{p+q-2, 2(p+q)-2, \ldots, N(p+q)-2\}$. The position $p-1$ is between the bands and $p-2$ is outside the bands or vice versa. Let S_1 be the component that contains $p-1$ and S_2 be the component that contains $p-2$. The partial word $u_{(2N,p,q)}$ is not $(2N, p, q)$-special of Type 2 since neither S_1 nor S_2 is two-isolated by Hole$(u_{(2N,p,q)})$. To see this, Definition 3(1) fails with $i = p-1$ or $i = p-2$. To show that $u_{(2N,p,q)}$ is not $(2N,p,q)$-special of Type 3, we can use Definition 4(1) with $i = p-1$ or $i = p-2$. To show that $u_{(2N,p,q)}$ is not $(2N, p, q)$-special of Type 1, we can use Definition 2(2) with $i = N(p+q) - 1 + q$ or

$i = N(p + q) — 2 + q.$

REFERENCES

1. N.J. Fine and H.S. Wilf, Uniqueness theorem for periodic functions, Proceedings of the American Mathematical Society 16, 109-114, (1965).
2. J. Berstel and L. Boasson, Partial words and a theorem of Fine and Wilf, Theoretical Computer Science 218, 135-141, (1999).
3. F. Blanchet-Sadri and R.A. Hegstrom, Partial words and a theorem of Fine and Wilf revisited, Theoretical Computer Science 270, 401-419, (2002).