Research Article
Compatibility and Conjugacy on Partial Arrays

S. Vijayachitra¹ and K. Sasikala²

¹Department of Science and Humanity, Sathyabama University, Chennai, India
²Department of Mathematics, St. Joseph’s College of Engineering, Chennai, India

Correspondence should be addressed to S. Vijayachitra; vijayachitra@live.com

Received 27 March 2016; Revised 28 July 2016; Accepted 3 August 2016

Academic Editor: Qi Dai

Copyright © 2016 S. Vijayachitra and K. Sasikala. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research in combinatorics on words goes back a century. Berstel and Boasson introduced the partial words in the context of gene comparison. Alignment of two genes can be viewed as a construction of two partial words that are said to be compatible. In this paper, we examine to which extent the fundamental properties of partial words such as compatibility and conjugacy remain true for partial arrays. This paper studies a relaxation of the compatibility relation called $k$-compatibility. It also studies $k$-conjugacy of partial arrays.

1. Introduction

The genetic information in almost all organisms is carried by molecules of DNA. A DNA molecule is a quite long but finite string of nucleotides of 4 possible types: $a$ (for adenine), $c$ (for cytosine), $g$ (for guanine), and $t$ (for thymine). The stimulus for recent works on combinatorics is the study of biological sequences such as DNA and protein that play an important role in molecular biology [1–3]. Sequence comparison is one of the primitive operations in molecular biology. Alignment of two sequences is to place one sequence above the other [2, 4] in order to make clear correspondence between similar letters or substrings of the sequences. Partial words appear in comparing genes. Indeed, alignment of two genes can be viewed as a construction of two partial words that are said to be compatible. In some cases, it allows insertion of letters which relate to errors or mismatches. A problem appears when the same gene is sequenced by two different labs that want to differentiate the gene expression. Also, when the same long sequence is typed twice into the computer, errors appear in typing.

Partial array $A$ of size $(m, n)$ over $\Sigma$, a finite alphabet, is partial function $A : \mathbb{Z}_+^2 \to \Sigma$, where $\mathbb{Z}_+$ is the set of all positive integers. In this paper, we extend the combinatorial properties of partial words to partial arrays. Also, this paper studies a relation called $k$-compatibility where a number of insertions and deletions are allowed as well as $k$-mismatches. The conjugacy result [6] which was proved for partial words is extended to partial arrays. $k$-Conjugacy of partial arrays is discussed.

2. Preliminaries on Partial Words

In this section, we give a brief overview of partial words [7].

**Definition 1.** Partial word $u$ of length $n$ over $A$, a nonempty finite alphabet, is partial map $u : \{1, 2, \ldots, n\} \to A$. If $1 \leq i \leq n$, then $i$ belongs to the domain of $u$ (denoted by Domain($u$)) in the case where $u(i)$ is defined, and $i$ belongs to the set of holes of $u$ (denoted by Hole($u$)), otherwise.

A word [8–10] is a partial word over $A$ with an empty set of holes.

**Definition 2.** Let $u$ be a partial word of length $n$ over $A$. The companion of $u$ (denoted by $u_\emptyset$) is map $u_\emptyset : \{1, 2, \ldots, n\} \to A \cup \{\emptyset\}$ defined by

$$u_\emptyset (i) = \begin{cases} u (i) & \text{if } i \in \text{Domain} (u) \\ \emptyset & \text{otherwise.} \end{cases}$$

(1)
The symbol \( \Diamond \) is viewed as a "do not know" symbol. Word \( u_0 = ba\Diamond ab\Diamond \) is the companion of the partial word. The length of the partial word is 6. \( D(u) = \{1, 2, 4, 5\} \). \( H(u) = \{3, 6\} \).

Let \( u \) and \( v \) be two partial words of length \( n \). Partial word \( u \) is said to be contained in partial word \( v \) (denoted by \( u \subset v \)), if Domain\( (u) \subset \text{Domain}(v) \) and \( u(i) = v(i) \) for all \( i \in \text{Domain}(u) \). Partial words \( u \) and \( v \) are called compatible (denoted by \( u \uparrow v \)), if there exists partial word \( w \) such that \( u \subset w \) and \( v \subset w \) (in which case we define \( u \uparrow v \) by \( u \uparrow v \) and \( v \uparrow v \) and Domain\( (u \uparrow v) = \text{Domain}(u) \cup \text{Domain}(v) \)). As an example, \( u_0 = aba\Diamond a \) and \( v_0 = ababa\Diamond a \).

The following rules are useful for computing with partial words:

(i) **Multiplication**: If \( u \uparrow v \) and \( x \uparrow y \), then \( ux \uparrow vy \).

(ii) **Simplification**: If \( ux \uparrow vy \) and \( |u| = |v| \), then \( u \uparrow v \) and \( x \uparrow y \).

(iii) **Weakening**: If \( u \uparrow v \) and \( w \subset u \), then \( w \uparrow v \).

**Lemma 3.** Let \( u, v, x, y \) be partial words such that \( ux \uparrow vy \).

(i) If \(|u| \geq |v|\), then there exist partial words \( w, z \) such that \( u = wz \), \( v \uparrow w \), and \( y \uparrow zw \).

(ii) If \(|u| \leq |v|\), then there exist partial words \( w, z \) such that \( b = wz \), \( v \uparrow w \), and \( x \uparrow zy \).

**Definition 4.** Two partial words \( u \) and \( v \) are called conjugate, if there exist partial words \( x \) and \( y \) such that \( u \subset xy \) and \( v \subset yx \).

**Definition 5.** Two partial words \( u \) and \( v \) are called \( k \)-conjugate, if there exist nonnegative integers \( k_1, k_2 \) whose sum is \( k \) and partial words \( x \) and \( y \) such that \( uc_{k_1}xy \) and \( vc_{k_2}yx \).

### 3. Preliminaries on Partial Arrays

This section is devoted to review the basic concepts on partial arrays [11].

**Definition 6.** Partial array \( A \) of size \((m, n)\) over \( \Sigma \), a nonempty set or an alphabet, is a partial function \( A : Z^2_+ \to \Sigma \), where \( Z_+ \) is the set of all positive integers. For \( 1 \leq i \leq m \), \( 1 \leq j \leq n \), and if \( A(i, j) \) is defined, then we say that \( (i, j) \) belongs to the domain of \( A \) (denoted by \( D(A) \)). Otherwise, we say that \( (i, j) \) belongs to the set of holes of \( A \) (denoted by \( H(A) \)).

An array [5] over \( \Sigma \) is a partial array over \( \Sigma \) with an empty set of holes.

**Definition 7.** If \( A \) is a partial array of size \((m, n)\) over \( \Sigma \), then the companion of \( A \) (denoted by \( A_\Diamond \)) is total function \( A_\Diamond : Z^2_+ \to \Sigma \cup \{\Diamond\} \) defined by

\[
A_\Diamond(i, j) = \begin{cases} A(i, j) & \text{if} \,(i, j) \in D(A) \\ \Diamond & \text{otherwise}, \end{cases}
\]

where \( \Diamond \notin \Sigma \).

**Example 8.** Partial array \( A = \begin{pmatrix} b & a & b \\ \Diamond & a & b \\ b & \Diamond & b \end{pmatrix} \) is the companion of partial array \( A \) of size \((3, 3)\), where \( D(A) = \{(1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 1), (3, 3)\} \), \( H(A) = \{(2, 1), (3, 2)\} \).

Let

\[
X = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},
\]

\[
Y = \begin{pmatrix} b_{11} & \cdots & b_{1n'} \\ \vdots & \ddots & \vdots \\ b_{m'1} & \cdots & b_{m'n'} \end{pmatrix}.
\]

By column catenation, we mean

\[
X \circ Y = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1n'} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_{m'1} & \cdots & b_{m'n'} \end{pmatrix}.
\]

By row catenation, we mean

\[
X \diamond Y = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \\ b_{11} & \cdots & b_{1n'} \\ \vdots & \ddots & \vdots \\ b_{m'1} & \cdots & b_{m'n'} \end{pmatrix}.
\]

If \( A \) and \( B \) are two partial arrays of equal size, then \( A \) is contained in \( B \) denoted by \( A \subset B \) if \( D(A) \subseteq D(B) \) and

\[
A(i, j) = B(i, j) \quad \forall (i, j) \in D(A).
\]

**Definition 9.** Partial arrays \( A \) and \( B \) are said to be compatible denoted by \( A \uparrow B \), if there exists partial array \( C \) such that \( A \subset C \) and \( B \subset C \).

### 4. Compatibility and \( k \)-Compatibility of Partial Arrays

#### 4.1. Compatibility

The rules mentioned for partial words are also true for partial arrays.

Let \( A, B, X, Y \) be partial arrays.

(i) **Multiplication**: If \( A \uparrow B \) and \( X \uparrow Y \), then \( AX \uparrow BY \) either by column catenation or by row catenation.

(ii) **Simplification**: If \( AX \uparrow BY \) either by column catenation or by row catenation with \( A \) and \( B \) being of same size, then \( A \uparrow B \) and \( X \uparrow Y \).
(iii) Weakening: If $A \uparrow B$ and $C \subset A$, then $C \uparrow B$.

Lemma 10. Let $A$, $B$, $X$, $Y$ be partial arrays such that $AX \uparrow BY$, either by column catenation or by row catenation.

(i) If order of $A \geq$ order of $B$, then there exist partial arrays $C$, $Z$ such that $A = CZ$, $B \uparrow C$, and $Y \uparrow ZX$.

(ii) If order of $A \leq$ order of $B$, then there exist partial arrays $C$, $Z$ such that $B = CZ$, $A \uparrow C$, and $X \uparrow ZY$.

4.2. $k$-Compatibility

Definition 11. If $A$ and $B$ are two partial arrays of same size and $k$ is nonnegative integer, then $A$ is said to be $k$-contained in $B$ denoted by $AC_k B$ if $D(A) \subset D(B)$ and there exists subset $E$ of $D(A)$ of cardinality $k$ called the error set such that

$$A(i,j) = B(i,j) \quad \forall (i,j) \in D(A) \setminus E,$$

$$A(i,j) \neq B(i,j) \quad \forall (i,j) \in E.$$  \hfill (7)

Definition 12. If $A$ and $B$ are two partial arrays of same order and $k$ is a nonnegative integer, then $A$ and $B$ are called $k$-compatible denoted by $A \mid_k B$ if there exist partial array $Z$ and nonnegative integers $k_1$, $k_2$ such that

(i) $AC_{k_1} Z$ with error set $E_1$;

(ii) $BC_{k_2} Z$ with error set $E_2$;

(iii) $E_1 \cap E_2 = \emptyset$;

(iv) $k_1 + k_2 = k$.

Example 13. $A = (b\ 0\ 0\ a \ b\ a\ b\ a\ b)$, $B = (b\ a\ b\ 0\ 0\ a\ b\ b\ b)$, then there exists partial array $Z = (b\ a\ b\ 0\ 0\ a\ b\ b\ b)$ with $E_1 = \{(1,1), (1,2), (2,1), (2,2), (2,3)\}$ and $k_1 = 2$, $k_2 = 1 \Rightarrow k = 3$; that is, $A \mid_k B$.

Equivalently, $A$ and $B$ are $k$-compatible, if there exists subset $E$ of $D(A) \cap D(B)$ of cardinality $k$ called the error set such that

(i) $A(i,j) = B(i,j) \forall (i,j) \in D(A) \cap D(B) \setminus E$;

(ii) $A(i,j) \neq B(i,j) \forall (i,j) \in E$.

If $A$ and $B$ are arrays, then $A \mid_k B$ means $A = B$. We sometimes use notation $A \mid_k B$, if set $E$ has cardinality $\leq k$.

Multiplication. If $A \mid_k B$ and $X \mid_k Y$, then $AX \mid_{k_1 + k_2} BY$ where $A$, $B$, $X$, and $Y$ are partial arrays and $k_1$, $k_2$ are nonnegative integers, using column catenation.

Example 14. $A = (\begin{array}{ccc} 0 & a & a \\ b & b & 0 \\ a & a & b \end{array})$, $B = (\begin{array}{ccc} b & b & 0 \\ 0 & a & a \\ b & b & 0 \end{array})$, $X = (\begin{array}{ccc} b & 0 & a \\ 0 & a & b \\ a & b & b \end{array})$, $Y = (\begin{array}{ccc} a & b & 0 \\ b & a & 0 \\ a & b & 0 \end{array})$.

$AX \mid_{k_1 + k_2} BY$.

Simplification. If $AX \mid_k BY$ and order of $A$ is equal to order of $B$, then $A \mid_k B$ and $X \mid_k Y$ for some $k_1$, $k_2$, satisfying $k_1 + k_2 = k$.

Example 15. $A = (\begin{array}{ccc} 0 & a & a \\ b & b & 0 \\ a & a & b \end{array})$, $B = (\begin{array}{ccc} b & b & 0 \\ 0 & a & a \\ b & b & 0 \end{array})$, $X = (\begin{array}{ccc} b & 0 & a \\ 0 & a & b \\ a & b & b \end{array})$, $Y = (\begin{array}{ccc} a & b & 0 \\ b & a & 0 \\ a & b & 0 \end{array})$.

$AX \mid_k BY \Rightarrow A \mid_3 B$ and $X \mid_3 Y$ with $5 + 3 = 8$.

Weakening. If $A \mid_k B$ and $Z \subset A$, then $Z \mid_{\leq k} B$.

Example 16. $A = (\begin{array}{ccc} 0 & a & a \\ b & b & 0 \\ a & a & b \end{array})$, $B = (\begin{array}{ccc} b & b & 0 \\ 0 & a & a \\ b & b & 0 \end{array})$, $Z = (\begin{array}{ccc} a & b & 0 \\ b & b & 0 \\ b & a & 0 \end{array})$.

$Z \mid_{\leq 7} B$ with $k = 7$.

Theorem 17. Let $A$ and $B$ be partial arrays of orders $a \times b$ and $a \times c$, respectively. If there exist array $Z$ of order $a \times d$ and integers $k_1$, $k_2$, $m$, and $n$ such that $AC_{k_1} Z^m$ with error set $E_1$ and $BC_{k_2} Z^n$ with error set $E_2$, then there exist integers $p$ and $q$ such that $A^p \mid_{\leq k} B^q$ with

$$k = |\{(D(A)(a, |b|, p) \cap E_1(a, |c|, q)) \cup (D(B)(a, |c|, q) \cap E_1(a, |b|, p))\}|.$$  \hfill (8)

Moreover, if $E_1(a, |b|, n) \cap E_2(a, |c|, m) = \emptyset$, then $A^p \mid_k B^q$.

Proof. Let $A$ and $B$ be partial arrays of $a \times b$ and $a \times c$, respectively. Let array $Z$ of order $a \times d$ exist such that, by using column catenation, $AC_{k_1} Z^m$ and $BC_{k_2} Z^n$ for some integers $k_1$, $k_2$, $m$, and $n$. Let $E_1$ be the error set of cardinality $k_1$ such that $A(i,j) = Z^m(i,j)$ for all $(i,j) \in D(A) \setminus E_1$ and $A(i,j) \neq Z^m(i,j)$ for all $(i,j) \in E_1$ and $E_2$ be the error set of cardinality $k_2$ such that $B(i,j) = Z^n(i,j)$ for all $(i,j) \in D(B) \setminus E_2$ and $B(i,j) \neq Z^n(i,j)$ for all $(i,j) \in E_2$. We have $A_{mk_1} C_{mk_2} Z^m$ with error set $E_1(a, |b|, n)$ of cardinality $nk_1$ and $B^m c_{mk_2} Z^n$ with error set $E_2(a, |c|, m)$ of cardinality $mk_2$.

Let $(1,1) \leq (i,j) \leq (a, d^m)$ and $Z^m(i,j) = a$ for some letter $a$. There are 4 possibilities.

Case 1. If $(i,j) \notin E_1(a, |b|, n)$ and $(i,j) \notin E_2(a, |c|, m)$, then $A^0(i,j) \in \{0, a\}$ and $B^0(i,j) \in \{0, a\}$. It does not give any error, when we align $A^0$ with $B^0$.

Case 2. If $(i,j) \notin E_1(a, |b|, n)$ and $(i,j) \in E_2(a, |c|, m)$, then $A^0(i,j) \in \{0, a\}$ and $B^0(i,j) = b$ for some $b \neq a$. It gives an error in the alignment of $A^0$ with $B^0$ only when $B^0(i,j) = a$ or when $(i,j) \in D(A)(a, |b|, n)$.

Case 3. If $(i,j) \in E_1(a, |b|, n)$ and $(i,j) \in E_2(a, |c|, m)$, then $B^0(i,j) \in \{0, a\}$ and $A^0(i,j) = b$ for some $b \neq a$. It gives an error in the alignment of $A^0$ with $B^0$ only when $B^0(i,j) = a$ or when $(i,j) \in D(B)(a, |c|, m)$.

Case 4. If $(i,j) \in E_1(a, |b|, n)$ and $(i,j) \in E_2(a, |c|, m)$, then $A^0(i,j) = b$ for some $b \neq a$ and $B^0(i,j) = c$ for some $c \neq a$. It gives an error in the alignment of $A^0$ with $B^0$ only when $B^0(i,j) = c$.

Therefore, if $E_1(a, |b|, n) \cap E_2(a, |c|, m) = \emptyset$ then $A^0 \mid_{\leq k} B^0$ with $k = |\{(D(A)(a, |b|, n) \cap E_2(a, |c|, m)) \cup (D(B)(a, |c|, m) \cap E_1(a, |b|, n))\}|$ and $E_1(a, |b|, n) \cap E_2(a, |c|, m) \neq \emptyset$ then $A^0 \mid_{\leq k} B^0$.  \hfill \square

Example 18. $A = (\begin{array}{ccc} a & b & 0 \\ b & b & 0 \\ b & b & 0 \end{array})$, $B = (\begin{array}{ccc} a & b & 0 \\ b & b & 0 \\ 0 & 0 & 0 \end{array})$, $Z = (\begin{array}{ccc} a & b & 0 \end{array})$. 


We have $A \subseteq Z_3$ with error set $E_1 = \{(1,2),(2,2),(2,3),(3,3)\}$, and $B \subseteq Z_2$ with error set $E_2 = \{(1,2),(2,2)\}$.

$k = 6$:

(i) $D(A) = \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,3)\}$.

$D(B) = \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,2),(3,3)\}$.

(ii) $D(A)(a,|b|,2) = \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,2),(3,3)\}$.

$D(B)(a,|c|,3) = \{(1,1),(1,2),(2,1),(2,2),(2,3),(3,2),(3,3)\}$.

(iii) $E_1(a,|b|,2) = \{(1,2),(2,2),(2,3),(3,3),(1,5),(1,6),(2,5)\}$.

$E_2(a,|c|,3) = \{(1,2),(2,2),(1,4),(2,4),(1,6),(2,6)\}$.

$k = \|D(A)(a,|b|,2) \cap E_2(a,|c|,3)\| = \|D(B)(a,|c|,3) \cap E_1(a,|b|,2)\| = 1$.

$k^2 = \|D(A)(a,|b|,2) \cap E_2(a,|c|,3)\| = \|D(B)(a,|c|,3) \cap E_1(a,|b|,2)\| = 1$.

$k = 9$: $A^2 \uparrow B^3 (A^2 \uparrow_a B^3)$. 

5. Conjugacy and $k$-Conjugacy of Partial Arrays

5.1. Conjugacy

Definition 19. Two partial arrays $A$ and $B$ of same order are called conjugate if there exist partial arrays $X$ and $Y$ such that $A \subseteq XY$ and $B \subseteq YX$ using row catenation or column catenation.

0-conjugacy on partial arrays with same order is trivially reflexive and symmetric but not transitive.

Example 20. $A = \begin{pmatrix} a & 0 & b \\ b & c & a \\ a & 0 & a \\ a & a & b \\ a & b & c \end{pmatrix}, B = \begin{pmatrix} b & c & 0 \\ a & 0 & b \\ a & a & 0 \\ a & b & 0 \end{pmatrix}, C = \begin{pmatrix} a & a & 0 \\ a & b & 0 \\ b & a & 0 \end{pmatrix}$.

By taking $X = (a \ 0 \ b)$ and $Y = (b \ c \ a)$, we get $A \subseteq XY$ and $B \subseteq YX$ showing that $A$ and $B$ are conjugate similarly and, by taking $X' = (b \ c \ 0)$ and $Y' = (a \ c \ b)$, we get $B \subseteq X'Y'$ and $C \subseteq Y'X'$ showing that $B$ and $C$ are conjugate. But $A$ and $C$ are not conjugate.

That is, conjugate relation is not transitive.

Proposition 21. Let $A$ and $B$ be nonempty partial arrays of same size. If $A$ and $B$ are conjugate, then there exists partial array $C$ such that $AC \uparrow CB$, either by column catenation or by row catenation.

Proof. Let $A$ and $B$ be nonempty partial arrays of same order. Suppose $A$ and $B$ are conjugate and let $X$, $Y$ be partial arrays such that $A \subseteq XY$ and $B \subseteq YX$ either by column catenation or by row catenation; then $AX \subseteq XYX$ and $XB \subseteq YXY$. So, for $C = X$, we have $AC \uparrow CB$. □

5.2. $k$-Conjugacy

Definition 22. Two partial arrays $A$ and $B$ of same order are $k$-conjugate, if there exist nonnegative integers $k_1, k_2$ whose sum is $k$ and partial arrays $X$ and $Y$ such that $AC_{k_1}XY$ and $BC_{k_2}YX$ with row or column catenation.

Theorem 23. Let $A$ and $B$ be nonempty partial arrays of same order. If $A$ and $B$ are $k$-conjugate, then there exists partial array $Z$ such that $AZ \uparrow_{\leq k} ZB$.

Proof. Let $A$, $B$ be two partial arrays of same order. Supposing that $A$ and $B$ are $k$-conjugate, then, by definition, there exist nonnegative integers $k_1$, $k_2$ whose sum is $k$ and partial arrays $X$ and $Y$ such that $AC_{k_1}XY$ with error set $E_1$ and $BC_{k_2}YX$ with error set $E_2$ using row catenation or column catenation accordingly.

Then, $AXC_{k_1}X$ with error set $E_1$ and $XB_{k_2}Y$ with error set $E_2$ is $\subseteq X$ such that $AC \uparrow_{\leq k} ZB$.

Example 24. Given $A = \begin{pmatrix} a & 0 & b \\ b & c & a \\ a & 0 & b \\ a & b & c \end{pmatrix}$.

There exist $X = (a \ 0 \ b)$ and $Y = (b \ a \ 0)$ with $AC_{\leq 3}XY$ and $Bc_{\leq 2}Y$, $k_1 + k_2 = 5$.

There exist $Z = (a \ 0 \ b)$ such that $AZ \uparrow_{\leq 5} ZB$.

6. Conclusion

Motivated by compatibility and conjugacy properties of partial words, we define the conjugacy of partial array and derive the compatibility properties of partial arrays. By giving relaxation to the compatibility relation, we discuss $k$-compatibility and $k$-conjugacy of partial arrays. We prove that, given partial arrays $A$, $B$ and integers $p, q$ satisfying $|A|^p = |B|^q$, we find $k$ such that $A^{p} \uparrow_{k} B^{q}$. Also, there exists partial array $Z$ such that $AZ \uparrow_{\leq k} ZB$.

Disclosure

S. Vijayachitra is a Research Scholar at Department of Science and Humanity Sathyabama University, Chennai, India.

Competing Interests

The authors declare that they have no competing interests.
References


