Hindusthan International Journal of Multidisciplinary Research (HIJOMR)
Volume 1 Issue 2, December 2024, pp 12-20

GRAPH RECURRENCE OF SPLICING SYSTEM

Dr. S. Anuradha[1], R. Anusha[2]

[1]Professor & Head,
Department of Mathematics, Data Science & Analytics
Hindusthan College of Arts & Science, Coimbatore
anuradha.s@hicas.ac.in

[2]Assistant Professor, Department of Mathematics,
Hindusthan College of Arts & Science, Coimbatore
anusha.r@hicas.ac.in

ABSTRACT

A graph recurrence sequence $x_0, x_1, x_2, \ldots\ldots, x_{2k-1}$ of vectors for splicing system were identified by the recurrence$x_{m+1} = A_n x_m$. The motivation of defining the graph recurrence sequences are to authenticate whether the given spliced graph is determined by a set of vectors and the existence of equivalence of two spliced graphs. The graph splicing isomorphism is analyzed for finite and infinite recurrence sequence of vectors. The recurrence relation is constructed for splicing system and its analytical solution had been evaluated for all positive integers.

KEYWORDS: *Splicing System, Graph Recurrence Sequence, Graph Splicing Isomorphism, Recurrence Relation*

1. INTRODUCTION

Tom Head [1] invented the notion of a splicing system in 1987 in an attempt to model certain biochemical reactions involving cutting and reconnecting strands of double-sided DNA. Since then the theory of splicing systems has become a new interesting area of Formal Languages Theory. DNA exists not only as Linear Molecules but also as Circular molecules. Investigating the Complex structures of genes, Rudolf Freund [3] suggested graphs as more suitable objects for modeling such graphs and exhibit the relation between regular graph splicing systems on graphs and splicing systems on strings. The graph splicing scheme of Freund with semi graphs studied by E. Sampath Kumar [8] to get a characterization of spliced semi graph to show some graph properties.Since Leonard Adleman [9] demonstrated that the recombinant used to solve combinatorial search problem, the field of DNA computing has developed very fast. It provides parallelism and high density storage when compared to silicon based computer. DNA sequences are three-dimensional objects in a three dimensional space. Hence graphs seem to be more suitable objects for describing complex three-dimensional objects independently from their actual position in three dimensional spaces. A DNA single-strand consists of four different types of units called nucleotides or bases strung together by an oriented backbone like beads on a wire. The bases are Adenine (A), Guanine (G), Cytosine (C), Thymine (T) and A can chemically bind to an opposing T on another single-strand while C can similarly bind to G. To encode information using DNA, one can choose an encoding scheme mapping the original alphabet onto over {A, T, C, G} and proceed to synthesize the obtained information-encoding strings as DNA single strands. The DNA strands representing the output of the computation can be read out and decoded. On this basis here we adopt graph an important tool for DNA-computational method.In this paper, we investigate the graph recurrence sequences from spliced graphs at the time of splicing DNA molecule in order to determine whether any two spliced graphs in terms of their adjacency matrices are isomorphic or not. The recurrence relation for spliced graphs had also been framed here.

2. PRELIMINARIES

2.1. Splicing System

Splicing is a model of the recombinant behavior of double-stranded molecules of DNA under the action of restriction enzymes and ligases. A single stranded of DNA is an oriented sequence of nucleotides A, C, G & T but since A can bind to T & G to C, two strands of DNA bind together to form a double stranded DNA molecule, if they have matching pairs of nucleotides when reading the second one along the reverse orientation.

2.2. Graph

A graph over V is a triple (N, E, L) where N is the set of nodes, E is the set of edges of the form (n, m) with n, m $\in$ N, n $\neq$ m and L is the function from N to V assigning a label from V to each node of N. The set of all graphs over V is denoted by γ (V).

2.3. Graph splicing scheme

A graph splicing scheme σ is a pair (V, P), where V is an alphabet & P is a finite set of graph splicing rules of the form ((h [1], E' [1])… (h [k], E' [k]); E), where $k \geq 1$& for all i with $1 \leq i \leq k$,

- h[i] $\in\gamma$ (V), h[i] = (N [i], E [i], L[i]),
- E'[i] $\subseteq$ E [i],
- The node sets N[i] are mutually disjoint, and E must obey the following rules:

1. Each edge (n, m) $\in$ E' [i] divided is supposed to be into two parts, i.e., the start part (n, m] and the end part [n, m).
2. The elements of E are of the form ((n, m], [n', m')), where (n, m) and (n', m') are edges from $U_{1 \leq i \leq k}$ E' [i].
3. Every element from {(n, m], [n, m) / (n, m) $\in U_{1 \leq i \leq k}$ E' [i]} must appear exactly once in a pair of E.

2.4. Graph splicing rule

Let P = ((h[1], E'[1),…, (h[k], E'[k]); E) be a graph splicing rule, and let $\gamma \in \gamma_c(V)$. If we can select k different graphs g [1]… g[k] from γ, then we can apply p to γ, which yields some $s \in \gamma_c(V)$ in the following way:

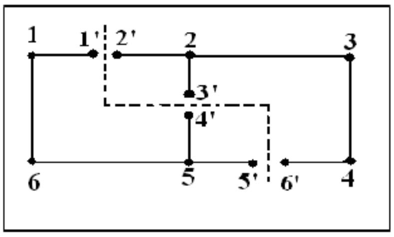
- For all i with $1 \leq i \leq k$, h [i] is a sub graph of g[i], where f[i] establishes the injective node embedding of h[i] into g[i].
- The union of g[1] … g[k] can be looked at as a single graph g in γ (V) and the union of the functions f[i] as a single function f embedding the h[i] into g. From g we now eliminate all edges from $U_{1 \leq i \leq k}$ f(E'[i]), but add all edges (f(n), f(m')) such that ((n, m], [n', m')) $\in$ E, which yields the uniquely determined union of k' weakly connected graphs g' [1], …, g' [k'].

2.5. 'n-cut' spliced graph

The graph obtained by splicing DNA molecule at ‘n-strands’ is called an ‘n-cut’ spliced graph.

Example 2.1

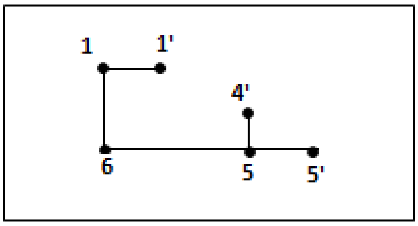
The following graph *G* represents ‘1-cut’ splicing (u_{C_1}) with the vertices {1', 2', 3', 4', 5', 6'} and the edges {(1, 1'), (2, 2'), (2, 3'), (5, 5'), (5, 4'), (4, 6')}. Here |N’| = 6 and |E’| = 6.

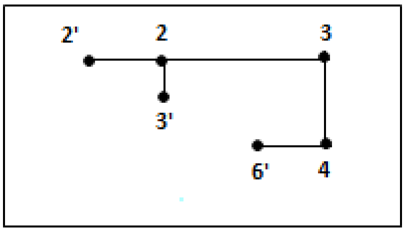


G

Figure 1: Representation of ‘1-cut’ splicing

At the time of ‘1-cut’ splicing, the graph *G* is splitted into the following two graphs G_1 and G_2.



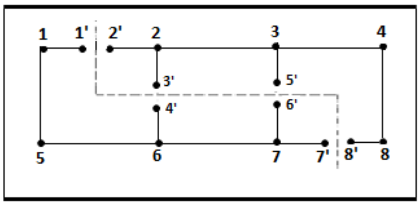


G_1 G_2

Figure 2: ‘1-cut’splicing split-ted graphs

Example 2.2:

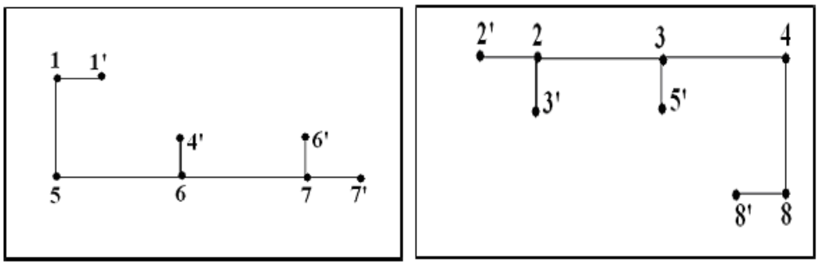
Let σ = ({A, T, C, G}, P), where P= {(Fig.2.5, Fig. 2.6{(1, 1'), (2', 2), (2, 3'), (4', 6), (6, 7), (7, 6'), (7, 7'),
(8', 8)})}. Here A, T, C and G denote the four deoxyribonucleotides that incorporate adenine, cytosine, guanine, thymine and with the enzyme the edge is splitted into two edges which undergoes splicing operation.



H

Figure 3: Representation of ‘2-cut’ splicing

The graph H at ‘2-cut’ splicing, we get



H_1 H_2

Figure 4:'2-cut'splicing splitted graphs

3. GRAPH RECURRENCE SEQUENCE OF VECTORS FOR SPLICING SYSTEM

A graph recurrence sequence $x_0, x_1, x_2, \dots\dots, x_{2k-1}$ of vectors for an 'n-cut' spliced graph G_n with $n \geq 1$ having $2k$ vertices $\forall k \geq 3$ is defined by the recurrence $x_{m+1} = A_n x_m$ for $m=0,1,2,\dots\dots,2n+2$ where A_n is an adjacency matrix of G_n and x_0 is an initial vector in R^{2k}. Each vector in the above sequence can be treated as a vertex labeling of G_n. The label at a given vertex at step $m+1$ obtained by summing the values at the adjacent vertices at step m.

Consider the initial vector of '1-cut' spliced graph G_1 as x_0= (1,0,0,0,0,0). Then the successive vectors will be

x_1= (0,1,0,0,0,0)
x_2= (1,0,1,0,0,0)
x_3= (0,2,0,1,0,0)
x_4= (2,0,3,0,1,1)
x_5= (0,5,0,5,0,0).

Similarly for '2-cut' spliced graph G_2, consider the initial vector as x_0= (1,0,0,0,0,0,0,0) and the successive vectors are x_1= (0,1,0,0,0,0,0,0)
x_2= (1,0,1,0,0,0,0,0)
x_3= (0,2,0,1,0,0,0,0)
x_4= (2,0,3,0,1,1,0,0)
x_5= (0,5,0,5,0,0,0,0)
x_6= (5,0,10,0,5,5,0,0)
x_7= (0,15,0,5,20,0,0,5,5).

Proceeding like this, we get the graph recurrence sequence of vectors for 'n-cut' spliced graph.

Let us assume the graph recurrence sequence of vectors as $X_1, X_2, X_3, \dots\dots$
$X_1 = x_0, x_1, x_2, x_3, x_4, x_5$
$X_2 = x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7$
$X_3 = x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9$ and so on.

From above, we observe that $X_1 \subset X_2, X_2 \subset X_3$,.....

Therefore, the graph recurrence sequence of vectors for 'n-cut' spliced graph is a subset of the graph recurrence sequence of vectors for '(n+1)-cut'spliced graph for all $n \geq 1$.

Result 3.1.

The variation of the cardinality of graph recurrence sequence of vectors of every consecutive cut spliced graph is two.

Result 3.2.

If the standard vector e_i is an initial vector of graph recurrence sequence for 'n-cut' spliced graph with 1 at coordinate i and the remaining $2k$-1 terms as 0. Then the graph recurrence sequence for 'n-cut' spliced graph is centered at vertex i. In the above example, the graph recurrence sequence of vectors $X_1, X_2, X_3, \dots\dots$ are centered at vertex 1. For an adjacency matrix A_n, G_{A_n} denotes the corresponding 'n-cut' spliced graph. B_n is the adjacency matrix of an 'n_1-cut' spliced graph G_{B_n} which is obtained by the rearrangement of the last $2k$-2 vertices of an 'n-cut' spliced graph G_{A_n}. In this case, G_{B_n} is said to have the same X_n-sequence as G_{A_n}. Also, the graph recurrence sequences for G_{A_n} and G_{B_n} are identical for all initial values of X_n. Moreover, there exist an Edmonds matrix E such that $A_n^m x = (E^{-1}B_nE)^m x$ $\forall\, m \geq 0$, $n \geq 1$ and $x \in X_n$.

3.1. Definition

A set $X_n \subset R^{2k}$ is said to distinguish an 'n-cut' spliced graph G_{A_n} from an 'n_1-cut' spliced graph G_{B_n} if G_{B_n} does not have the same X_n-sequence as G_{A_n}. (i.e) when the rearrangement of all the vertices of G_{A_n} occurs, X_n distinguishes G_{A_n} from G_{B_n}.

3.2. Definition

A set $X_n \subset R^{2k}$ is said to determine an 'n-cut' spliced graph G_{A_n} if X_n distinguishes 'n-cut' spliced graph G_{A_n} from any graph not isomorphic to G_{A_n}. Since X_n distinguishes 'n-cut' spliced graph G_{A_n} from an '(n+1)-cut'spliced graph $G_{A_{n+1}}$ which is not isomorphic to G_{A_n}, we say that X_n determines G_{A_n}.

Theorem 3.1.

Let A_n and B_n be $2k \times 2k$ matrices $\forall\, k \geq 3$ and $x \in R^{2k}$. If $A_n^m x = B_n^m x$ for $0 \leq m \leq 2k$-1 then $A_n^m x = B_n^m x$ for all $m \geq 0$.

Proof. Let us prove the theorem by induction on m. We have, $A_n^m x = B_n^m x$ for $0 \leq m \leq 2k$-1. If s_n is the degree of the minimal polynomial for B_n then $A_n^m x = B_n^m x$ for $0 \leq m \leq s_n$. Consider $A_n^m x = B_n^m x$ for $0 \leq m \leq r_n$. We get, $B_n^{r_n} = p(B_n)$ where $deg(p) < s_n$ (since after reducing by the minimal polynomial). Now, $A_n^{r_n+1}x = A_n A_n^{r_n} x = A_n B_n^{r_n} x = A_n p(B_n)x = A_n p(A_n)x = B_n p(B_n)x = B_n^{r_n+1}x$. Also, $deg[xp(x)] \le s_n$. Therefore, $A_n^m x = B_n^m x$ for all $m \ge 0$.

Theorem 3.2.

Let G_{A_n} be an 'n-cut' spliced graph and $X_n \subset R^{2k}$. If $\{A_n^m x/x \in X_n,\ m \ge 0\}$ spans R^{2k} then G_{A_n} is not determined by X_n.

Proof. Let us assume $A_n^m x = (E^{-1}B_nE)^m x\ \ \forall m \ge 0,\ x \in X_n$ and E is an Edmonds matrix. Then $A_n(A_n^m x) = A_n^{m+1}x = (E^{-1}B_nE)^{m+1}x = (E^{-1}B_nE)\,(E^{-1}B_n^m E)x = (E^{-1}B_nE)\,A_n^m x$. We have $A_n^m x = (E^{-1}B_n^m E)x\ \ \forall x \in R^{2k}$ (since $\{A_n^m x/x \in X_n,\ m \ge 0\}$ spans R^{2k}). Therefore, $A_n^m = (E^{-1}B_n^m E)$. (i.e) A_n^m and B_n^m are identical matrices. So G_{A_n} and G_{B_n} are isomorphic to each other. Hence G_{A_n} is not determined by X_n.

4. GRAPH SPLICING ISOMORPHISM

A permutation of a vertex set $[2k]\ \forall k \ge 3$ on a vector $v = (v_0, v_1, v_2, \ldots\ldots, v_{2k-1})$ is defined as a bijection from $[2k]$ to itself. (i.e) A bijection map $f\colon [2k] \to [2k]$ (where $[2k] = \{0,1,\ldots,2k\text{-}1\}$) on a vector $v = (v_0, v_1, v_2, \ldots\ldots, v_{2k-1})$ is defined by $f(2k) = (v_{f(0)}, v_{f(1)}, v_{f(2)}, \ldots\ldots, v_{f(2k-1)})$.

Consider X_n as the sequence of vectors of 'n-cut' spliced graphs G_{A_n}. Let Y_n be the sequence of vectors obtained by reordering the last $2k\text{-}2$ vertices of 'n_1-cut' spliced graphs G_{B_n}. Two sequences X_n and Y_n are said to be equivalent ($X_n \equiv Y_n$) when $X_n = (x_0, x_1, x_2, \ldots, x_{2k-1})\ \ \forall\ k \ge 3$ & $Y_n = (y_0, y, y_2, \ldots\ldots, y_{2k-1})\ \forall k \ge 3$ and if there exist a single permutation $f(x_i) = y_i\ \forall\ 0 \le i \le 2k\text{-}1$.

The 'n-cut' & 'n_1-cut' spliced graphs G_{A_n} and G_{B_n} on $2k$ vertices are said to be equivalent ($G_{A_n} \equiv G_{B_n}$) if there exist a bijection $h : [2k] \to [2k]$ such that $A_{n_i} \equiv B_{n_{h(i)}}\ \ \forall\, i \in [2k]$ where A_{n_i} & $B_{n_{h(i)}}$ are the finite sequences of vectors on G_{A_n} and G_{B_n}. The finite sequence of vector A_{n_i} is given by $\{A_1^0 e_i, A_2^1 e_i, \ldots, A_n^{2k-1} e_i\}$ for the graph recurrence centered at vertex i. From the above argument, we obtain two cases.

***Case i*:** A bijectionbetween the finite vertex sets of 'n-cut' & 'n_1-cut' spliced graphs G_{A_n} & G_{B_n} such that the graph finite recurrence sequences centered at corresponding vertices are the same. Therefore, there is a "**fake isomorphism**" between G_{A_n} & G_{B_n}.

***Case ii*:** We have, A_{n_i} is a finite sequence of vector of an 'n-cut' spliced graph G_{A_n}. Theorem 3.1 shows the equivalence of the corresponding infinite recurrence sequence of vectors. Hence, obviously there is an "**isomorphism**" between G_{A_n} & G_{B_n} and also they are equivalent.

5. RECURRENCE RELATION OF 'N-CUT' SPLICED GRAPH

For the following splicing system, F_n denotes the number of cuts in the splicing system after 'n' times. Therefore, $F_1 = 1$.

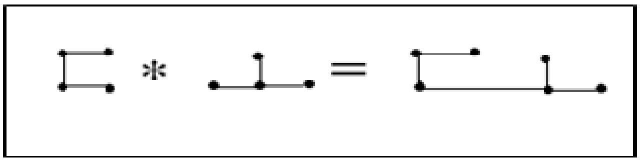
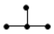


Figure 5: $F_0 + C = F_1$

From figure 5, F_1 is a concatenation(*) of F_0 and C.

Here  is denoted by C and F_0=0 i.e. $F_0 + C = F_1$.

Similarly, for '2-cut' spliced graph

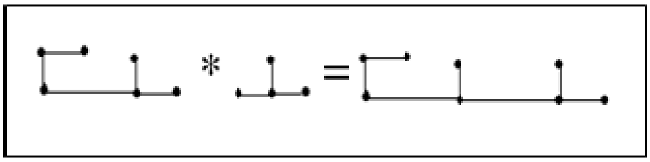


Figure 6: $F_1 + C = F_2$

From figure 6, F_2 is a concatenation(*) of F_1 and C.

i.e. $F_1 + C = F_2$.

Proceeding like this, we get $F_{n-1} + C = F_n$.

Therefore, $F_n - F_{n-1} = C$ is the recurrence relation of 'n-cut' spliced graph.

Consider the value of C as the number of edges in C. (i.e.) C=3.

Hence $F_n - F_{n-1} = 3 \; \forall \, n \geq 1$.

Solving the above recurrence relation, we get

$F_n = F_{n-1} + 3$. Now replace n by $n-1$, $F_{n-1} = F_{n-2} + 3$.

Substitute F_{n-1} in to F_n, $F_n = (F_{n-2} + 3) + 3 \Rightarrow F_n = F_{n-2} + 2.3$. By replacing n by $n-2$, we get $F_{n-2} = F_{n-3} + 3$. Substitute F_{n-2} in to F_n to obtain $F_n = F_{n-2} + 2.3 \Rightarrow F_n = (F_{n-3} + 3) + 2.3 \Rightarrow F_n = F_{n-3} + 3.3$. In general, $F_n = F_{n-k} + k.3$ where $k=1, 2, 3, n-1$. Substitute $k=n-1$ to obtain, $F_n = F_{n-(n-1)} + (n-1).3 \Rightarrow F_n = F_1 + (n-1).3$

Substitute the initial condition F_1=1 in the above equation, we get $F_n = 1 + (n-1).3 \Rightarrow F_n = 3n - 2$ as the required solution.

6. CONCLUSION

In this paper, we discussed the graph recurrence sequence of vectors to validate whether the given spliced graphs are determined and isomorphic. It is hoped that this approach will contribute immensely for the development of DNA computing.

REFERENCES

[1] T. Head, Formal Language Theory and DNA. An analysis of the generative capacity of recombinant behaviors. Bulletin of Mathematical Biology, 49, 1987, 735-759.

[2] J. Dassow, V. Mitrana, Self Cross-over Systems, Discrete Mathematics and Theoretical Computer Science, Springer Series, G.Paun (Ed.), 1998, 283-294.

[3] R. Freund, Splicing Systems on Graphs, In Proceedings on Intelligence in Neural and Biological Systems, IEEE Press, 1995, 189-194.

[4] K. Gnanamalar David, K.G. Subramanian and D. Gnanaraj Thomas, A Note on Graph Splicing Languages, *Lecture Notes in Computer Science*, Vol.2340, Springer-Verlag, 381-390, 2001.

[5] F. Harary, *Graph Theory*, Addison-Wesley, Reading, Mass, (1969).

[6] N. Jonoska, 3D DNA Patterns and Computation, *Proc. Molecular Computing*, India, 1998, 20-32

[7] S. JeyaBharathi, J. Padmashree, S. Sinthanaiselvi, "*Semigraph structure on DNA splicing system*", Sixth international conference on Bio-inspired computing: Theories and applications, IEEE-computer society, Penang, Malaysia, 2011, 182-187.

[8] E. Sampathkumar, Semigraphs and their Applications, *Report on the DST project, 2000.*

[9] L. Adleman, Molecular computation of solutions to combinatorial problems, Science, 266(1994), 1021-24.

[10] S. JeyaBharathi, K. Thiagarajan, M. Kameshwari, "*DNA Splicing System through semi graph*", Proc. International Conference on Emerging Trends in Mathematics and Computer Applications, India, 2010, 75.

[11] A. Vince, Graph recurrence, European Journal of Combinatorics 24(2003) 15-32.

Hindusthan International Journal of Multidisciplinary Research (HIJOMR)
Volume 1 Issue 2, December 2024, pp 12-20

AUTHOR 1 PROFILE

Author – 1

Dr.S.Anuradha, Professor, Head Department of Mathematics, Data Science and Analytics is having 29 Years Teaching and Research Experience. She received Best Teacher Award 2013 in Akshaya Institute of Management studies Coimbatore , Grade A award for Team Leader in Task Based Training (TBT-2019) conducted by e-Yantra Lab Setup Initiative(eLSI), IIT BOMBAY, Vocational Service Award 2021 from Rotary Club of Coimbatore DownTown in 2021, Most Accomplished Spoken Tutorial Educator Resource Award in 2023 by SPOKEN TUTORIAL, IIT BOMBAY. She is a Teacher invitee in 34th Mid Year Meeting of Indian Academy of Sciences in the year 2023. She published 73 International Journal, 12 Books, 9 Book Chapter and 1 patent. She is Editorial Board Member in International Research Conference Series-Scientific and Technical Committee and Editorial board Member on mathematical and Computational Sciences and Reviewer in International Journals of American Journal of Applied Scientific Research, Science Publishing Group, Advances in Applied Research, www.IndianJournals.com,Latin American Applied Research, Applied Mathematics & Information Sciences, Nature Publishing Group. 7 PhD Scholars and 18 MPhil Scholars awarded Degree under her guidance.

Author – 2

Mrs.R.Anusha is currently working as an Assistant Professor at the Department of Mathematics, Hindusthan College of Arts & Science, Coimbatore. In 2011, she completed her Master of Philosophy in Madurai Kamaraj University. She has presented three papers in International Conferences. She has published three research papers in International Journals. Also, she published two books in 2024. Her areas of interest include Graph Theory, Statistics and Differential Equations.