A SUFFICIENT CONDITION FOR HANNA NEUMANN PROPERTY OF SUBMONOIDS OF A FREE MONOID

SHUBH NARAYAN SINGH AND K. V. KRISHNA

ABSTRACT. Using automata-theoretic approach, Giambruno and Restivo have investigated on the intersection of two finitely generated submonoids of the free monoid over a finite alphabet. In particular, they have obtained Hanna Neumann property for a special class of submonoids generated by finite prefix sets. This work continues their work and provides a sufficient condition for Hanna Neumann property for the entire class of submonoids generated by finite prefix sets. In this connection, a general rank formula for the submonoids which are accepted by semi-flower automata is also obtained.

INTRODUCTION

Howson proved that the intersection of two finitely generated subgroups of a free group is finitely generated [7]. In 1956, Hanna Neumann improved the result that if $H$ and $K$ are finite rank subgroups of a free group, then

$$\widetilde{rk}(H \cap K) \leq 2\widetilde{rk}(H)\widetilde{rk}(K),$$

where $\widetilde{rk}(N) = \max(0, rk(N) - 1)$ for a subgroup $N$ of rank $rk(N)$. Further, Neumann conjectured that

$$\widetilde{rk}(H \cap K) \leq \widetilde{rk}(H)\widetilde{rk}(K),$$

which is known as Hanna Neumann conjecture [12]. In 1990, Walter Neumann proposed a stronger form of the conjecture called strengthened Hanna Neumann conjecture (SHNC) [13]. Meakin and Weil proved SHNC for the class of positively generated subgroups of a free group [9]. The conjecture has recently been settled by Mineyev (cf. [10, 11]) and announced independently by Friedman (cf. [3, 4]).

In contrast, it is not always true that the intersection of two finitely generated submonoids of a free monoid is finitely generated. It appears that the intersection problem for submonoids of free monoids is much more complex than the analogous problem for subgroups of free groups. In particular, Hanna Neumann property for submonoids of a free monoid is of special interest. Two finitely generated submonoids $H$ and $K$ of a free monoid are said to satisfy Hanna Neumann property (in short, HNP), if $H$ and $K$ satisfy the inequality $(*)$. There are several contributions in the literature to study the intersection of two submonoids of a free monoid.

In 1972, Tilson proved that the intersection of free submonoids of the free monoid over a finite alphabet is free [16]. In connection to HNP, Karhumäki obtained a result for submonoids of rank two of the free monoid over a finite alphabet. In fact,
Karhumäki proved that the intersection of two submonoids of rank two is generated either by a set of at most two words or by a regular language of a special form \([8]\). Using automata-theoretic approach, Giambruno and Restivo have investigated on the rank and HNP of certain submonoids of a free monoid \([6]\).

In \([6]\), Giambruno and Restivo introduced the concept called semi-flower automata (in short, SFA). An SFA accepts a finitely generated submonoid of the free monoid over the underlying alphabet, and vice versa. Moreover, if an SFA is deterministic, it accepts the submonoid generated by a finite prefix set. Conversely, the submonoid generated by a finite prefix set is accepted by a deterministic SFA with at most one ‘branch point going in’ (in short, bpi). Using SFA, Giambruno and Restivo have initiated the investigations on the intersection of two submonoids generated by finite prefix sets. It is clear that the product automaton of two deterministic SFA accepts the intersection of the submonoids accepted by the deterministic SFA. If the product automaton is also semi-flower, then clearly the intersection is also finitely generated. Giambruno and Restivo have considered two submonoids which are accepted by deterministic SFA with a unique bpi such that the product automaton is also semi-flower with at most one bpi. They have shown that such submonoids satisfy HNP. Further, if the product automaton has more than one bpi, they have provided some examples of submonoids which fail to satisfy HNP.

Recently in \([14, 15]\), Singh and Krishna have obtained a condition on the product automaton with two bpi’s, so that the respective submonoids satisfy HNP.

The present work is in the direction of addressing HNP for the entire class of submonoids generated by finite prefix sets. This work generalizes the work of Singh and Krishna and provides a sufficient condition for HNP of two submonoids generated by finite prefix sets. Further, a general rank formula for the submonoids which are accepted by SFA is also obtained. The main work of the paper is presented in Section 2. Before that, in Section 1, the necessary preliminary concepts and results are presented. The paper is concluded in Section 3.

### 1. Preliminaries

In this section, we present the necessary background material from \([1, 5, 6]\). Let \(A\) be a finite set called an alphabet with its elements as letters. The free monoid over \(A\) is denoted by \(A^*\) and \(\varepsilon\) denotes the empty word – the identity element of \(A^*\). It is known that every submonoid of \(A^*\) is generated by a unique minimal set of generators. Thus, if \(H\) is a submonoid of \(A^*\), then the rank of \(H\), denoted by \(rk(H)\), is defined as the cardinality of the minimal set of generators \(X\) of \(H\), i.e. \(rk(H) = |X|\). Further, the reduced rank of \(H\) is defined as \(\bar{rk}(H) = \max(0, rk(H) - 1)\) and it is denoted by \(\bar{rk}(H)\). A set of words, say \(X\), is said to be a prefix set if no element of \(X\) is a proper prefix of another element in \(X\).

Let \(A\) be an alphabet. An automaton over \(A\) is a quadruple \((Q, I, T, F)\), where \(Q\) is a finite set called the set of states, \(I\) and \(T\) are subsets of \(Q\) called the sets of initial and final states, respectively, and \(F \subseteq Q \times A \times Q\) called the set of transitions. Clearly, by denoting the states as vertices/nodes and the transitions as labeled arcs, an automaton can be represented by a digraph in which initial and final states shall be distinguished appropriately.

In what follows, let \(A = (Q, I, T, F)\) be an automaton over \(A\). A state, say \(q\), of \(A\) is called a branch point going in, in short bpi, if the number of transitions coming into \(q\) (i.e. the indegree of \(q\) – the number of arcs coming into \(q\) – in the digraph of
The transition relation can be seen as a partial function from $Q$ to itself, i.e. the set of states whose outdegree – the number of arcs going out of the state – in the digraph of $A$ is $i$.

A path in $A$ is a finite sequence of consecutive arcs in its digraph. For $p_i \in Q$ ($0 \leq i \leq k$) and $a_j \in A$ ($1 \leq j \leq k$), let

$$p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} p_2 \xrightarrow{a_3} \cdots \xrightarrow{a_k} p_{k-1} \xrightarrow{a_k} p_k$$

be a path, say $P$, in $A$ that is starting at $p_0$ and ending at $p_k$. In this case, we write $i(P) = p_0$ and $f(P) = p_k$. The word $a_1 \cdots a_k \in A^*$ is the label of the path $P$. A null path is a path from a state to itself labeled by $\varepsilon$.

A path in $A$ is called simple if all the states on the path are distinct. A path that starts and ends at the same state is called as a cycle, if it is not a null path. A cycle with all its states are distinct is called a simple cycle. Other notions related to paths, viz. subpath, prefix path and suffix path, can be defined in a usual way or one may refer to [6].

The language accepted by $A$, denoted by $L(A)$, is the set of words in $A^*$ that are the labels of the paths from an initial state to a final state. A state, say $q$, of $A$ is accessible (respectively, coaccessible) if there is a path from an initial state to $q$ (respectively, a path from $q$ to a final state). The trim part of $A$, denoted by $A^T$, is the automaton obtained from $A$ by considering only the accessible and coaccessible states, and the respective transitions between them. Note that $L(A) = L(A^T)$. If $A = A^T$, then we say $A$ is trim. An automaton is deterministic if it has a unique initial state and there is at most one transition defined for a state and a letter, i.e. the transition relation can be seen as a partial function from $Q \times A$ to $Q$.

An automaton is called a monoidal automaton if it is trim with a unique initial state that is equal to a unique final state. Further, a monoidal automaton is called a semi-flower automaton, in short SFA, if all the cycles in the automaton visit the unique initial-final state.

If $A = (Q, I, T, F)$ is a monoidal automaton, we denote the initial-final state by $q_0$. In which case, we simply write $A = (Q, q_0, q_0, F)$. Further, if $A$ is an SFA, let us denote by $C_A$ the set of simple cycles (passing through $q_0$) in $A$ and by $Y_A$ the set of their labels.

In the following theorem we state the correspondence between SFA and finitely generated submonoids of a free monoid.

**Theorem 1.1** ([5]). If $A$ is an SFA over $A$, then $Y_A$ is finite and $A$ accepts the submonoid generated by $Y_A$ in $A^*$. Moreover, if $A$ is deterministic, then $Y_A$ is a prefix set and it is the minimal set of generators of the submonoid accepted by $A$. Conversely, let $X$ be a finite subset of $A^*$ and let $H$ be the submonoid generated by $X$; then there exists an SFA accepting $H$. Furthermore, if $X$ is a prefix set, then there exists a deterministic SFA with at most one bpi accepting $H$.

We now present the concept of product automaton and discuss the state of the art of the intersection problem of two submonoids generated by finite prefix sets.

Let $A_1 = (Q_1, I_1, T_1, F_1)$ and $A_2 = (Q_2, I_2, T_2, F_2)$ be two automata defined over $A$. The product automaton $A_1 \times A_2$ is the automaton

$$(Q_1 \times Q_2, I_1 \times I_2, T_1 \times T_2, F)$$
over the alphabet \( A \) such that
\[
((p, p'), a, (q, q')) \in \mathcal{F} \iff (p, a, q) \in \mathcal{F}_1 \text{ and } (p', a, q') \in \mathcal{F}_2
\]
for all \( p, q \in Q_1, p', q' \in Q_2 \) and \( a \in A \).

Notice that, if \( A_1 \) and \( A_2 \) are deterministic, then so is \( A_1 \times A_2 \). But, if \( A_1 \) and \( A_2 \) are trim, then \( A_1 \times A_2 \) need not be trim. Further,
\[
L(A_1 \times A_2) = L((A_1 \times A_2)^T) = L(A_1) \cap L(A_2).
\]

Let \( H \) and \( K \) be submonoids generated by finite prefix sets of words over \( A \). In view of Theorem [11] suppose \( A_H \) and \( A_K \) are deterministic SFA over \( A \) with at most one bpi accepting \( H \) and \( K \), respectively. Clearly, \( (A_H \times A_K)^T \) is a deterministic monoidal automaton accepting \( H \cap K \). In order to consider the case that \( H \cap K \) is finitely generated, one could restrict \( (A_H \times A_K)^T \) to be semi-flower.

With this hypothesis, we discuss on HNP of \( H \) and \( K \) as follows.

**Case 1.** \( A_H \) or \( A_K \) has no bpi’s: In this case, \( (A_H \times A_K)^T \) has no bpi’s. Consequently, \( H \cap K \) is cyclic, so that \( H \) and \( K \) satisfy HNP.

**Case 2.** \( A_H \) and \( A_K \) have unique bpi: In this case, \( (A_H \times A_K)^T \) can have arbitrary number of bpi’s. Thus, the problem is considered into various subcases and we only know the following.

I. \( (A_H \times A_K)^T \) has at most one bpi: In this subcase, if \( (A_H \times A_K)^T \) is semi-flower, then \( H \) and \( K \) satisfy HNP (cf. [6, Theorem 3.6]).

II. \( (A_H \times A_K)^T \) has two bpi’s: In this subcase, even if \( (A_H \times A_K)^T \) is semi-flower, \( H \) and \( K \) need not satisfy HNP (cf. [15, Example 3.7]). However, if \( (A_H \times A_K)^T \) is an SFA with two bpi’s having a unique path from one bpi to the other, then \( H \) and \( K \) satisfy HNP (cf. [15, Corollary 3.11]).

In general, if \( (A_H \times A_K)^T \) has more than one bpi, there are several examples of \( H \) and \( K \) which fail to satisfy HNP (cf. [5, 6]). Thus, if \( (A_H \times A_K)^T \) has arbitrary number of bpi’s, in this work we would investigate on certain conditions so that \( H \) and \( K \) satisfy HNP. For that purpose, we would require the following supplementary results from [6].

**Proposition 1.2.** Let \( A \) be an alphabet of cardinality \( n \). If \( A = (Q, q_0, q_0, \mathcal{F}) \) is a deterministic SFA over \( A \), then
\[
|\mathcal{F}| - |Q| = \sum_{i=2}^{n} |BPO_i(A)|(i - 1).
\]

**Proposition 1.3.** Let \( A \) be an alphabet of cardinality \( n \) and let \( A_1 \) and \( A_2 \) be two deterministic automata over \( A \). If \( c_i = |BPO_i(A_1)| \) and \( d_i = |BPO_i(A_2)| \), for each \( i \in \{1, \ldots, n\} \), then
\[
|BPO_i(A_1 \times A_2)| \leq \sum_{t \leq r, s \leq n} c_t d_s.
\]

**Proposition 1.4.** Let \( \langle c_1, \ldots, c_n \rangle \) and \( \langle d_1, \ldots, d_n \rangle \) be two finite sequences of natural numbers; then
\[
\sum_{t=2}^{n} (t - 1) \left( \sum_{t \leq r, s \leq n} c_r d_s \right) \leq \left( \sum_{i=2}^{n} (i - 1)c_i \right) \left( \sum_{j=2}^{n} (j - 1)d_j \right).
\]
2. Main Results

In this section we generalize the work of Singh and Krishna in [15]. First we obtain the rank of the submonoid of a free monoid that is accepted by an SFA. Then we proceed to obtain a condition for HNP of two submonoids of a free monoid which are generated by finite prefix sets.

2.1. BPR and Rank. We begin with introducing a concise notation for an SFA in which only the initial-final state, bpi’s and the respective paths between them will be considered along with their labels. We call this as bpi’s and paths representation, in short BPR, of the semi-flower automaton.

**Definition 2.1.** Let \( A = (Q, q_0, q_0, F) \) be an SFA over \( A \); the BPR of \( A \) is a quadruple \( A' = (Q', q_0, q_0, F') \), where

(i) \( Q' = BPI(A) \cup \{q_0\} \), and

(ii) \( F' \) is the finite subset of \( Q' \times A^* \times Q' \) defined by \( (p_0 = p, x, q = p_k) \in F' \) if and only if there exist distinct \( p_1, \ldots, p_{k-1} \in Q \setminus Q' \) and \( x = a_1 \cdots a_k \), for \( a_i \in A \), such that \( (p_{i-1}, a_i, p_i) \in F \) for all \( 1 \leq i \leq k \), i.e.

\[
p = p_0 \xrightarrow{a_1} p_1 \xrightarrow{a_2} p_2 \xrightarrow{a_3} \cdots \xrightarrow{a_{k-1}} p_{k-1} \xrightarrow{a_k} p_k = q
\]

is a simple path from \( p \) to \( q \) (or simple cycle, when \( p = q \)) labeled by \( x \) in which the intermediate nodes, if any, are outside \( Q' \).

**Remark 2.2.** By adopting the digraph representation of an automaton, we can draw a digraph for the BPR of an SFA. Here, the arcs are labeled by the labels (words) of respective simple paths (or simple cycles) of the SFA.

**Example 2.3.** The BPR of the SFA given in **Figure 1** is shown in **Figure 2**. Here, we distinguish the initial-final state \( q_0 \) by double circles.

**Remark 2.4.** Let \( A = (Q, q_0, q_0, F) \) be an SFA and let \( A' = (Q', q_0, q_0, F') \) be the BPR of \( A \).

(i) Every cycle in \( A' \) passes through the state \( q_0 \).

(ii) The number of simple cycles in \( A \) is equal to the number of simple cycles in \( A' \).
(iii) For any $p, q \in Q'$, the number simple paths from $p$ to $q$ in $A$ is equal to the number of simple paths from $p$ to $q$ in $A'$.

**Proposition 2.5.** Let $A$ be an SFA and let $A'$ be the BPR of $A$. There is a linear ordering $\preceq$ on the states of $A'$ such that

(i) the initial-final state $q_0$ is the least element, and
(ii) for $j \neq 0$, if $q_j \preceq q_i$, then there is no arc from $q_j$ to $q_i$ in $A'$.

**Proof.** Construct the digraph $G$ from the digraph of $A'$ by removing the arcs which are leaving out of the initial-final state $q_0$. By Remark 2.4(i), $G$ is a directed acyclic graph. Define a relation $\leq$ on the nodes of $G$ by

$p \leq q$ if and only if there is a simple path from $q$ to $p$ in $G$.

As the null path is a simple path from a node to itself in $G$, clearly $\leq$ is reflexive.

Since there are no cycles in $G$, it can be observed that $\leq$ is anti-symmetric and transitive. Thus, the relation $\leq$ is a partial ordering on the nodes of $G$. Since every state in $A$ is coaccessible, the initial-final state $q_0$ is the least element with respect to $\leq$.

As every partial ordering can be extended to a linear ordering, consider a linear ordering $\preceq$ of the nodes of $G$ which is an extension of $\leq$. Thus, the obtained linear ordering $\preceq$ is the desired one. □

**Remark 2.6.** By applying a topological sort algorithm (e.g., refer [2]) on the directed acyclic graph $G$, one can get a linear ordering as described in Proposition 2.5.

In what follows, by a topological ordering of the bpi’s of an SFA is meant a linear ordering on the states (possibly, except the initial-final state) of its BPR as in Proposition 2.5.

**Example 2.7.** Let $A$ be the SFA given in Figure 1. A topological ordering of the bpi’s of $A$ is

$p_2, p_4, p_3, p_1$.

Notice that there is no arc from $p_3$ or $p_4$ to $p_1$ in the BPR of $A$ (cf. Figure 2). Hence, in a topological ordering of the bpi’s of $A$, the bpi $p_1$ can come at any
position after $p_2$. Thus, the possible other topological orderings are $p_2, p_1, p_4, p_3$ and $p_2, p_4, p_1, p_3$.

**Notation 2.8.** Let $\mathcal{A}$ be an SFA and let $\mathcal{A}'$ be the BPR of $\mathcal{A}$. If we say that $\mathcal{A}$ has $m$ bpi's, we always assume that $q_1, q_2, \ldots, q_m$ are the bpi's of $\mathcal{A}$, which are considered in a topological ordering. That is, 

$$q_1 \preceq q_2 \preceq \cdots \preceq q_m.$$ 

As per the ordering, we also fix the following numbers in the BPR $\mathcal{A}'$.

(i) For $1 \leq i \leq m$, $\kappa_i$ always refer to the number of arcs from the state $q_0$ to the bpi $q_i$ in $\mathcal{A}'$.

(ii) For $1 \leq i, j \leq m$, $\kappa_{ij}$ always refer to the number of arcs from the bpi $q_i$ to the bpi $q_j$ in $\mathcal{A}'$.

**Remark 2.9.** As per the topological ordering of the bpi's of $\mathcal{A}$, we have $\kappa_{ij} = 0$ for all $j \geq i > 1$. If the initial-final state $q_0$ of $\mathcal{A}$ is a bpi, then clearly $q_1 = q_0$ so that, for $j \geq 1$, $\kappa_{1j} = \kappa_j$; otherwise, $\kappa_{1j} = 0$.

**Remark 2.10.** In the digraph of $\mathcal{A}$ (as well as in $\mathcal{A}'$), the indegree of a bpi $q_j$, $1 \leq j \leq m$, is given by the expression

$$\kappa_j + \sum_{i=j+1}^{m} \kappa_{ij}.$$ 

The following lemma is useful in obtaining the rank of an SFA.

**Lemma 2.11.** Let $\mathcal{A}$ be an SFA and let $p$ be the first bpi in a topological ordering of the bpi's of $\mathcal{A}$, i.e. $p \preceq q$, for all bpi's $q$; then

(i) there is a unique simple path from $p$ to the initial-final state $q_0$, and

(ii) every cycle in $\mathcal{A}$ visits $p$.

**Proof.** If the initial-final state $q_0$ is a bpi, then clearly $p = q_0$. In which case, the null path is the unique simple path from $p$ to $q_0$. And, since $\mathcal{A}$ is semi-flower, every cycle in $\mathcal{A}$ visits $p$. If $q_0$ is not a bpi, we proceed as follows.

(i) Since $p$ is coaccessible, there is a path from $p$ to the state $q_0$. Now suppose there are two different paths $P_1$ and $P_2$ with labels $u$ and $v$, respectively, from $p$ to the state $q_0$. Let $w$ be the label of longest suffix path $P'$ which is in common between the paths $P_1$ and $P_2$. As the state $q_0$ is not a bpi, $w \neq \varepsilon$. But then $i(P')$ will be a bpi different from $p$. This a contradiction to the choice of $p$. Thus, there is a unique simple path from $p$ to $q_0$.

(ii) Suppose there is a cycle that is not visiting $p$. Then the cycle contributes one to the indegree of the state $q_0$. Also, from above (i), there is a path from $p$ to the state $q_0$. This implies that the state $q_0$ is a bpi; a contradiction.

**Corollary 2.12.** Let $\mathcal{A}$ be an SFA. If $p$ is the first bpi in a topological ordering of the bpi's of $\mathcal{A}$, then $p$ is the first bpi in any topological ordering of the bpi's of $\mathcal{A}$.

Let $\mathcal{A}$ be an SFA. Now we are ready to present our first main result of the paper on the rank of the submonoid $L(\mathcal{A})$. The rank of $L(\mathcal{A})$ can be characterized using the bpi's of $\mathcal{A}$. First note that, if there is no bpi in $\mathcal{A}$, then clearly the rank of $L(\mathcal{A})$ is either 0 or 1. If $\mathcal{A}$ has at least one bpi, we have the following theorem.
Theorem 2.13. Let $\mathcal{A}$ be an SFA and $m \geq 1$. If $\mathcal{A}$ has $m$ bpi’s, then

$$rk(L(\mathcal{A})) \leq \sum_{i=1}^{m} \kappa_{i} \overline{\kappa_{i0}},$$

(1)

where $\overline{\kappa_{i0}}$ is the number of simple paths from the bpi $q_{i}$ to the initial-final state $q_{0}$. The number $\overline{\kappa_{i0}}$ can be given by the recursive formula

$$\overline{\kappa_{10}} = 1 \text{ and } \overline{\kappa_{i0}} = \sum_{j=1}^{i-1} \kappa_{ij} \overline{\kappa_{j0}}, \text{ for } i > 1.$$

Moreover, if $\mathcal{A}$ is deterministic, then the equality holds in (1).

Proof. Let $q_{1} \preceq q_{2} \preceq \cdots \preceq q_{m}$ be the bpi’s of $\mathcal{A}$. It is known from Theorem 1.1 that

$$rk(L(\mathcal{A})) \leq |Y_{\mathcal{A}}| \leq |C_{\mathcal{A}}|.$$

We prove the result by showing that $|C_{\mathcal{A}}|$, the number of simple cycles in $\mathcal{A}$ passing through the state $q_{0}$, is equal to the righthand side of (1), i.e. we show that

$$|C_{\mathcal{A}}| = \sum_{i=1}^{m} \kappa_{i} \overline{\kappa_{i0}}.$$

By Remark 2.4(ii), $|C_{\mathcal{A}}| = |C_{\mathcal{A}'}|$, where $C_{\mathcal{A}'}$ is the number of simple cycles in the BPR $\mathcal{A}'$ of $\mathcal{A}$.

For $1 \leq i \leq m$, let $\nu_{i}$ be the number of simple cycles in $\mathcal{A}'$ that are passing through the bpi $q_{i}$ but not through any bpi $q_{j}$ with $j > i$. Clearly,

$$|C_{\mathcal{A}'}| = \sum_{i=1}^{m} \nu_{i}.$$

We conclude the result by arguing that $\nu_{i} = \kappa_{i} \overline{\kappa_{i0}}$, for $1 \leq i \leq m$.

In case $i = 1$, $\nu_{1}$ is the number of simple cycles in $\mathcal{A}'$ that are passing through the bpi $q_{1}$ but not through any other bpi. First note that, by Lemma 2.11 there is unique simple path from $q_{1}$ to $q_{0}$ so that $\overline{\kappa_{10}} = 1$. Now, each simple cycle that is counted in $\nu_{1}$ is merely an arc from $q_{0}$ to $q_{1}$ followed by the unique simple path from $q_{1}$ to $q_{0}$. Thus, the number of simple cycles counted in $\nu_{1}$ is the number of arcs from $q_{0}$ to $q_{1}$, i.e. $\kappa_{1}$. Hence, we have

$$\nu_{1} = \kappa_{1} = \kappa_{1} \overline{\kappa_{10}}.$$

For $i > 1$, as per the topological ordering, $\nu_{i}$ is clearly obtained by multiplying the number of arcs from $q_{0}$ to the bpi $q_{i}$ and the number of simple paths from $q_{i}$ to $q_{0}$. That is,

$$\nu_{i} = \kappa_{i} \overline{\kappa_{i0}}$$

as desired. Now, we obtain the recursive formula for $\overline{\kappa_{i0}}$. For $1 \leq t < i$, let $\mu_{it}$ be the number of simple paths in $\mathcal{A}'$ from the bpi $q_{i}$ to $q_{0}$ that are passing through the bpi $q_{t}$ but not through any other bpi $q_{j}$ with $j > t$. Clearly, $\overline{\kappa_{i0}} = \sum_{t=1}^{i-1} \mu_{it}$. But, for $1 \leq t < i$, the number $\mu_{it}$ is nothing else but the product of the number of arcs
from \( q_i \) to \( q_t \) and the number of simple paths from \( q_t \) to \( q_0 \), i.e. \( \mu_{it} = \kappa_{it} \kappa_{t0} \). Hence, we have the recursive formula

\[
\kappa_{i0} = \sum_{t=1}^{i-1} \kappa_{it} \kappa_{t0}.
\]

If \( \mathcal{A} \) is deterministic, then by Theorem 1.1 we have

\[
\text{rk}(L(\mathcal{A})) = |C_A| = \sum_{i=1}^{m} \kappa_{i} \kappa_{i0}.
\]

Now, Theorem 2.10 of [6] is an immediate corollary as stated below. We will use this corollary in one of our main results.

**Corollary 2.14.** If \( \mathcal{A} = (Q, q_0, q_0, F) \) is an SFA with a unique bpi, then

\[
\text{rk}(L(\mathcal{A})) \leq \kappa_1 = |F| - |Q| + 1.
\]

Moreover, if \( \mathcal{A} \) is deterministic, then the equality holds.

**Example 2.15.** Let us consider the topological ordering

\[
p_2 \preceq p_4 \preceq p_3 \preceq p_1
\]

of the bpi's of the SFA \( \mathcal{A} \) given in FIGURE 1. That is, \( q_1 = p_2, q_2 = p_4, q_3 = p_3 \) and \( q_4 = p_1 \). Accordingly, \( \kappa_1 = 1, \kappa_2 = 1, \kappa_3 = 2 \) and \( \kappa_4 = 3 \). Also, \( \kappa_{41} = 1, \kappa_{42} = 0, \kappa_{43} = 0, \kappa_{31} = 1, \kappa_{32} = 2 \) and \( \kappa_{21} = 1 \). Since \( \mathcal{A} \) is deterministic, we have

\[
\text{rk}(L(\mathcal{A})) = 11.
\]

### 2.2. Hanna Neumann Property.

In this subsection, we obtain a sufficient condition for HNP of two submonoids which are accepted by deterministic SFA with a unique bpi. The following lemma is useful in obtaining the proposed result.

**Lemma 2.16.** Let \( \mathcal{A} = (Q, q_0, q_0, F) \) be an SFA and \( m \geq 1 \). If \( \mathcal{A} \) has \( m \) bpi’s, then

\[
|F| - |Q| + 1 \geq \text{rk}(L(\mathcal{A})) = \sum_{i=2}^{m} \left( (\kappa_{i} - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_{ij} - 1) \right) + \sum_{j=1}^{m} \kappa_{j} \kappa_{j0} - 1
\]

Moreover, if \( \mathcal{A} \) is deterministic, then the equality holds.

**Proof.** Since the number of transitions \( |F| \) of \( \mathcal{A} \) is the total indegree (i.e. the sum of indegrees of all the states) of the digraph of \( \mathcal{A} \), by Remark 2.10 we have

\[
|F| = |Q| - m + \sum_{j=1}^{m} \left( \kappa_{j} + \sum_{i=j+1}^{m} \kappa_{ij} \right).
\]
Thus,
\[ |\mathcal{F}| - |Q| + 1 = \kappa_1 + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} \]
\[ = \kappa_1 + \sum_{i=2}^{m} \kappa_i \kappa_{i0} + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \kappa_i \kappa_{i0}. \]

Consequently,
\[ |\mathcal{F}| - |Q| + 1 = \kappa_1 + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} \]
\[ = \kappa_1 + \sum_{i=2}^{m} \kappa_i \kappa_{i0} + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \kappa_i \kappa_{i0}. \]

Now, by Theorem 2.13 and simple algebraic manipulations, we have
\[ |\mathcal{F}| - |Q| + 1 \geq rk(L(A)) + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \kappa_i \kappa_{i0} \]
\[ = rk(L(A)) + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \kappa_i \kappa_{i0} \]
\[ = rk(L(A)) + \sum_{j=2}^{m} (\kappa_j - 1) + \sum_{j=1}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \kappa_i \kappa_{i0} \]
\[ = rk(L(A)) + \sum_{j=2}^{m} (\kappa_j - 1) - \sum_{i=2}^{m} \kappa_i \kappa_{i1} (\kappa_i - 1) + \sum_{j=2}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \sum_{j=2}^{m} \kappa_i \kappa_{ij} \kappa_{j0} \]
\[ = rk(L(A)) - \sum_{i=2}^{m} (\kappa_i - 1) (\kappa_i1 - 1) + \sum_{j=2}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \sum_{j=2}^{m} \kappa_i \kappa_{ij} \kappa_{j0} \]
\[ = rk(L(A)) - \sum_{i=2}^{m} (\kappa_i - 1) (\kappa_i1 - 1) + \sum_{j=2}^{m} \sum_{i=j+1}^{m} \kappa_{ij} - \sum_{i=2}^{m} \sum_{j=2}^{m} \kappa_i \kappa_{ij} \kappa_{j0}, \]
as \kappa_{ij} = 0 \text{ for all } j \geq i > 1
\[ = rk(L(A)) - \sum_{i=2}^{m} (\kappa_i - 1) (\kappa_i1 - 1) \sum_{i=2}^{m} \kappa_i \kappa_{i0} - 1 \]
\[ = rk(L(A)) - \sum_{i=2}^{m} (\kappa_i - 1) (\kappa_i1 - 1) \sum_{i=2}^{m} \kappa_i \kappa_{i0} - 1 \]
as \kappa_{ij} = 0 \text{ for all } j \geq i > 1.

Thus,
\[ |\mathcal{F}| - |Q| + 1 \geq rk(L(A)) - \sum_{i=2}^{m} (\kappa_i - 1)(\kappa_i1 - 1) + \sum_{j=2}^{m} \kappa_i \kappa_{i0} - 1 \]. □
Now, by Proposition 1.2 we have the following corollary.

**Corollary 2.17.** Let $A$ be an alphabet of cardinality $n$ and let $A$ be a deterministic SFA over $A$. For $m \geq 1$, if $A$ has $m$ bpi’s, then

$$rk(L(A)) = \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right)$$

$$+ \sum_{t=2}^{n} |BPO_t(A)|(t - 1) + 1.$$

**Theorem 2.18.** Let $A_H$ and $A_K$ be deterministic SFA over $A$ each with a unique bpi accepting submonoids $H$ and $K$, respectively. For $m \geq 1$, if the automaton $(A_H \times A_K)^T$ is an SFA with $m$ bpi’s, say $q_1, q_2, \ldots, q_m$ considered in a topological ordering, then

$$\tilde{rk}(H \cap K) \leq \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right) + \tilde{rk}(H)\tilde{rk}(K),$$

where $\kappa_i$ is the number of arcs from the initial-final state to $q_i$ and $\kappa_{ij}$ is the number of arcs from $q_i$ to $q_j$ in the BPR of $(A_H \times A_K)^T$.

**Proof.** Let $A$ be an alphabet of cardinality $n$. For $m \geq 1$, note that

$$\tilde{rk}(H \cap K) = rk(L(A_H \times A_K)) - 1$$

$$= \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right)$$

$$+ \sum_{t=2}^{n} |BPO_t(A_H \times A_K)|(t - 1) \text{ by Corollary 2.17}$$

$$\leq \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right)$$

$$+ \sum_{t=2}^{n} (t - 1) \left( \sum_{t\leq r, s \leq n} c_r d_s \right) \text{ by Proposition 1.3}$$

where $c_r = |BPO_r(A_H)|$ and $d_s = |BPO_s(A_K)|$. Further, by Proposition 1.4 we have
\[ \tilde{r}k(H \cap K) \leq \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right) \\
+ \left( \sum_{i=2}^{n} (i - 1)c_i \right) \left( \sum_{j=2}^{n} (j - 1)d_j \right) \\
= \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right) + \tilde{r}k(H)\tilde{r}k(K) \]
by Corollary 2.14 and Proposition 1.2.

Hence the result. \[ \square \]

We now state a sufficient condition for Hanna Neumann property of the submonoids under consideration.

**Corollary 2.19.** In addition to the hypothesis of Theorem 2.18, if there is no path between any two bpi’s \( q_i \) and \( q_j \), for \( i, j > 1 \), and there is a unique simple path from each bpi to \( q_1 \) in the automaton \( (A_H \times A_K)^T \), then
\[ \tilde{r}k(H \cap K) \leq \tilde{r}k(H)\tilde{r}k(K). \]

**Proof.** For \( i, j > 1 \), if there is no path between the bpi’s \( q_i \) and \( q_j \), then \( \kappa_{ij} = 0 \). Further, for \( i \geq 2 \), if there is a unique simple path from each bpi \( q_i \) to \( q_1 \), then the path cannot pass through any other bpi. Thus, we have \( \kappa_{i1} = 1 \) so that
\[ \sum_{i=2}^{m} \left( (\kappa_i - 1)(\kappa_{i1} - 1) + \sum_{j=2}^{i-1} \kappa_{ij}(\kappa_i \kappa_{j0} - 1) \right) = 0. \]
Hence, by Theorem 2.18
\[ \tilde{r}k(H \cap K) \leq \tilde{r}k(H)\tilde{r}k(K). \]
\[ \square \]

3. **Conclusion**

This work considers the intersection problem of two submonoids of a free monoid which are generated by finite prefix sets. In particular, this work has obtained a sufficient condition for Hanna Neumann property for the class of submonoids generated by finite prefix sets. In that connection, a general rank for the submonoids which are accepted by semi-flower automata is also obtained. Thus, this work addresses one of the problems, viz. the prefix case, posed by Giambruno and Restivo in the conclusions of the paper [6]. However, there is a lot more to investigate on the general problem concerning the intersection of two arbitrary submonoids of a free monoid. For instance, even in the prefix case, one could investigate on the necessary and sufficient conditions for Hanna Neumann property. On the other hand, the intersection problem of two submonoids generated by finite non-prefix sets of words is of particular interest. For this problem, the rank formula that is obtained (for nondeterministic automata) in this paper may be useful.
Acknowledgements

The authors are very much thankful to anonymous referees for their valuable
comments which improved the manuscript; particularly, for pointing out a mistake
in Theorem 2.13.

References

[1] J. Berstel and D. Perrin. Theory of Codes, volume 117 of Pure and Applied Mathematics.
Academic Press Inc., 1985.
[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. Introduction to Algorithms. MIT
Press, second edition, 2001.
[3] J. Friedman. Linear algebra and the Hanna Neumann Conjecture, Preprint, 2011.
[4] J. Friedman. Sheaves on graphs, their homological invariants, and a proof of the Hanna Neu-
mann conjecture, Preprint, 2011. arXiv:1105.0120v2.
[5] L. Giambruno. Automata-theoretic Methods in Free Monoids and Free Groups. PhD thesis,
Universit degli Studi di Palermo, Palermo, Italy, 2007.
[6] L. Giambruno and A. Restivo. An automata-theoretic approach to the study of the intersection
of two submonoids of a free monoid. RAIRO – Theor. Inform. Appl., 42(3): 503–524, 2008.
[7] A. G. Howson. On the intersection of finitely generated free groups. J. London Math. Soc., 29:
428–434, 1954.
[8] J. Karhumäki. A note on intersection of free submonoids of a free monoid. Semigroup Forum,
29: 183–205, 1984.
[9] J. Meakin and P. Weil. Subgroups of free groups: a contribution to the Hanna Neumann
conjecture, Geom. Dedicata, 94: 33–43, 2002.
[10] I. Mineyev. Groups, graphs, and the Hanna Neumann Conjecture, Preprint, 2011.
[11] I. Mineyev. Submultiplicativity and the Hanna Neumann Conjecture, Ann. of Math., 175:
393–414, 2012.
[12] H. Neumann. On the intersection of finitely generated subgroups of free groups. Publ. Math.
Debrecen, 4: 186–189, 1956.
[13] W. D. Neumann. On intersections of finitely generated subgroups of free groups, Lecture
Notes in Math., 1456: 161–170, 1990.
[14] S. N. Singh and K. V. Krishna. On the rank of the intersection of two submonoids of a free
monoid. Contributed talk at A³: Abstract Algebra and Algorithms Conference, Eger, Hungary,
August 14–17, 2011.
[15] S. N. Singh and K. V. Krishna. The rank and Hanna Neumann property of some submonoids
of a free monoid. Ann. Math. Inform., To appear, 2011.
[16] B. Tilson. The intersection of free submonoids of a free monoid is free. Semigroup Forum, 4:
345–350, 1972.

Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati, India
E-mail address: {shubh, kvk}@iitg.ac.in