On the Shannon Cover of Shifts of Finite Type

Daniel P. B. Chaves, Cecilio Pimentel, and Bartolomeu F. Uchôa-Filho

***Abstract*— A shift space is a collection of sequences of symbols from a finite alphabet satisfying certain constraints. A shift of finite type is a shift whose constraints can be represented by a finite list of forbidden blocks. Every shift of finite type can also be represented by a labeled directed graph that has the property that every bi-infinite walk on the graph generates an allowed sequence by reading off the labels of its edges. It is both of theoretical and practical interest to find the minimal graph (called the Shannon cover), i.e., the one with the fewest vertices, presenting a shift of finite type. The main contribution of this paper is an efficient iterative vertex-minimization algorithm that considers the higher edge graph as the initial graph.**

***Keywords*— Constrained sequences, labeled graphs, Shannon cover, symbolic dynamics, sofic shifts.**

I. Introduction

The discrete sequences used in several applications to transmit or storage digital information often have to satisfy restrictions on the occurrence of consecutive symbols [1]. The set of all bi-infinite sequences satisfying a certain constraint is referred to in the Symbolic Dynamics literature as a *shift space*. A shift space is called *shift of finite type* if it may be specified in terms of a finite list of forbidden strings. Every shift of finite type can also be specified by a labeled directed graph, called the *presentation* of the shift space [2, Theorem 3.1.5].

A finite directed graph (or simply a *graph*) G consists of a finite set of vertices $\mathcal{V}(G)$, and a finite set of directed edges $\mathcal{E}(G)$ connecting the vertices. Each edge e has an initial vertex $i(e)$ and a terminal vertex $t(e)$. A labeled graph is a pair $\mathcal{G} = (G, \mathcal{L})$, where G is the underlying graph of $\mathcal{G}$, and the labeling $\mathcal{L} : \mathcal{E}(G) \to \mathcal{A}$ assigns to each edge a symbol from the alphabet $\mathcal{A}$. A labeled graph $\mathcal{G}$ is *right-resolving* if, for each vertex $I \in \mathcal{V}(G)$, the outgoing edges of I are labeled distinctly. The *Shannon cover* is a right-resolving presentation with the smallest number of vertices.

The Shannon cover for a shift of finite type is obtained, in general, via the following procedure: First, find an initial right-resolving presentation, then apply the Moore algorithm (see, for example,[2, p. 92], [3, p. 1661]) to recursively identify all classes of equivalent vertices (in the sense described in the next section). If there are r such classes, the Shannon cover has r vertices (one vertex is identified with a class), and the procedure described in [2, p. 78] shows how to obtain the edges and the labels of the Shannon cover. Even though this vertex-minimization algorithm is applicable to any right-resolving presentation, we give a characterization of the initial presentation which yields a simpler procedure to construct the Shannon cover. Section II contains background material on Symbolic Dynamics. In Section III we describe several properties of an initial presentation based on a labeling on the so-called higher edge graph. Our main result (Theorem 1) and the vertex-minimization algorithm are presented in this section. Section IV summarizes the conclusions of this work.

II. Background on Shift Spaces, Graphs and Sofic Shifts

This section summarizes relevant background material from Symbolic Dynamics. A detailed discussion on the topics covered here is found in [2]. See [3] for further reference. Let $\mathcal{A}^{\mathbb{Z}}$ be the set of all bi-infinite sequences (called points) $x = \cdots x_{-2}x_{-1}x_0x_1\cdots$ of symbols drawn from a alphabet $\mathcal{A}$ of finite size, namely $|\mathcal{A}|$. A finite sequence of consecutive symbols from $\mathcal{A}$ is called a *block*. We use the notation $x_{[i,j]} = x_ix_{i+1}\cdots x_j$ to specify a block which occurs in the point x, starting at position i and ending at position j. We say that a point $x \in \mathcal{A}^{\mathbb{Z}}$ *contains* a block w of length $M > 0$ if there exists an index i such that $w = x_{[i,i+M-1]}$. Let $\mathcal{F}$ be a set of blocks over $\mathcal{A}$, called *set of forbidden blocks*. A shift space X is the subset of $\mathcal{A}^{\mathbb{Z}}$ consisting of all points not containing any block from $\mathcal{F}$. A shift space is called an *M-step shift of finite type* if the length of the longest block in $\mathcal{F}$ is $M+1$. The set of all blocks of length n which occur in points of a shift space X is denoted by

Daniel P. B. Chaves and Cecilio Pimentel is with the Communications Research Group - CODEC, Department of Electronics and Systems, Federal University of Pernambuco, P.O. Box 7800 - 50711-970 - Recife - PE, ddselec@ig.com.br, cecilio@ufpe.br.

Bartolomeu F. Uchôa-Filho is with the Communications Research Group - GPqCom, Department of Electrical Engineering, Federal University of Santa Catarina, 88040-900 - Florianópolis - SC, uchoa@eel.ufsc.br.

The authors acknowledge partial support of this research by the Brazilian National Council for Scientific and Technological Development (CNPq) under Grant 302402/2002-0 and 302568/2002-6.

$\mathcal{B}_n(X)$. The *language* of the shift-space X is the set

$$\mathcal{B}(X) = \bigcup_{n=0}^{\infty} \mathcal{B}_n(X).$$

Two shift spaces are equal if and only if they have the same language [2, Proposition 1.3.4]. We say that X is *irreducible* if for every pair of blocks $u, v \in \mathcal{B}(X)$, there is a block $w \in \mathcal{B}(X)$ such that the concatenation uwv is also in $\mathcal{B}(X)$.

Let $\Phi : \mathcal{B}_{m+n+1}(X) \to \mathcal{M}$ be a *block map* from allowed $(m+n+1)$-blocks in X to symbols in an alphabet $\mathcal{M}$. The *sliding block code* with *memory* m and *anticipation* n *induced* by Φ is the map $\phi : X \to \mathcal{M}^{\mathbb{Z}}$ that transform $x \in X$ into a new point $y \in \mathcal{M}^{\mathbb{Z}}$, where $y_i = \Phi(x_{[i-m,i+n]})$. The special case when $m = n = 0$ is called *1-block code*. If Y is a subshift of $\mathcal{M}^{\mathbb{Z}}$, and $\phi(X) \subset Y$, we write $\phi : X \to Y$. If ϕ is onto, it is called a *factor code from* X *onto* Y. The image of X under ϕ is a shift space [2, Theorem 1.5.13]. A sliding block code $\phi : X \to Y$ is a *conjugacy from* X *to* Y, if it is invertible. In such a case, X and Y are called *conjugate*. Define the $(M+1)$th *higher block code* over X, $\beta_{M+1} : X \to Y$ as

$$(\beta_{M+1}(x))_i = x_{[i,\ i+M]}.$$

It is a sliding block code with memory $m = 0$ and anticipation M. The image of X under β_{M+1} is called the $(M+1)$th *higher block shift* of X, denoted by $X^{[M+1]}$. Consider $\mathcal{B}_{M+1}(X)$ as an alphabet, the shifts X and $X^{[M+1]}$ are conjugate since the inverse of β_{M+1} is the 1-block code $\psi : X^{[M+1]} \to X$ induced by the block map $\Psi : \mathcal{B}_{M+1}(X) \to \mathcal{A}$ defined as

$$(\psi(y))_i = \Psi(a_i a_{i+1} \dots a_{i+M}) = a_i \in \mathcal{A}.$$

Lemma 1: Let $\phi : X \to Y$ be a factor code with memory m and anticipation n. If X is irreducible, then Y is irreducible.

Proof: By contradiction, suppose that X is irreducible and Y is reducible. Since ϕ is a factor code, there exist two points x and x' in X such that two blocks $\phi(x)_{[i_1,j_1]} = u$ and $\phi(x')_{[i_2,j_2]} = v$ in $\mathcal{B}(Y)$ can be chosen in such a way that there is no block $w \in \mathcal{B}(Y)$ so that $uwv \in \mathcal{B}(Y)$. Suppose that u and v are the images of the blocks $x_{[i_1-m,\ j_1+n]} = \ell$ and $x'_{[i_2-m,\ j_2+n]} = k$, where $\ell, k \in \mathcal{B}(X)$. Since X is irreducible, there is a point $x'' \in X$, $x'' = \dots\ell\dots k\dots$, but there is no $y'' \in Y$ so that $y'' = \phi(x'')$. If this point y'' existed, there would be $w \in \mathcal{B}(Y)$ so that $uwv \in \mathcal{B}(Y)$, which is a contradiction. Therefore, if X is irreducible then Y must be irreducible. ■

Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph. We write $I \xrightarrow{a} J$ if there is an edge in G from I to J labeled a. A *path* in G is a block of edges $\pi = e_1 e_2 \cdots e_n$ such that the terminal vertex of e_i is the initial vertex of e_{i+1}. The label of π is the block $\mathcal{L}(\pi) = \mathcal{L}(e_1)\mathcal{L}(e_2)\cdots\mathcal{L}(e_n)$. A graph G is irreducible if for every pair of vertices $I, J \in \mathcal{V}(G)$ there is a path from I to J. A vertex $I \in \mathcal{V}(G)$ is *stranded* if either no edges start at I or no edges terminate at I. A graph is *essential* if it has no stranded vertices. A labeled graph is irreducible (or essential) if its underlying graph is irreducible (or essential).

A *walk* on G is a bi-infinite sequence of edges $\xi = \cdots e_{-1} e_0 e_1 \cdots$ such that $t(e_i) = i(e_{i+1})$, for all i. A *sofic shift* $X_{\mathcal{G}}$ is the set of sequences obtained by reading off the labels of walks on G. We say that $\mathcal{G}$ presents $X_{\mathcal{G}}$, or $\mathcal{G}$ is a presentation of $X_{\mathcal{G}}$. All sofic shifts are shift spaces [2, Theorem 3.1.4] and every shift of finite type is sofic [2, Theorem 3.1.5]. A block $w \in \mathcal{B}(X_{\mathcal{G}})$ is said to be generated by a path π in G if $w = \mathcal{L}(\pi)$. The incoming or outgoing edges of a stranded vertex cannot possibly occur in any walk on the graph. Then, if $\mathcal{G}$ presents $X_{\mathcal{G}}$, $\mathcal{G}$ has a unique essential labeled subgraph that presents $X_{\mathcal{G}}$. The follower set of a vertex $I \in \mathcal{G}$, denoted by $F_{\mathcal{G}}(I)$, is the set of all blocks of all lengths that can be generated by paths in G starting from I. The set of all such blocks of length n is denoted by $F^n_{\mathcal{G}}(I)$. Two vertices I and J are called *equivalent* if they have the same follower set. A presentation is follower-separated if $F_{\mathcal{G}}(I) = F_{\mathcal{G}}(J)$ implies that $I = J$. A block $w \in \mathcal{B}(X_{\mathcal{G}})$ is a *synchronizing block* for $\mathcal{G}$ if all paths in G that generate w terminate at the same vertex, say I. In this case, w focuses to I.

III. The Graph Presentation of Shifts of Finite Type

In this section, we will demonstrate the main results necessary to the determination of the Shannon cover of an M-step shift of finite type X. Without loss of generality, we will assume that the forbidden blocks in $\mathcal{F}$ all have the same length $M+1$. We will denote by $\mathcal{G}_{X^{[M+1]}} = (G_{X^{[M+1]}}, \mathcal{L}_X)$ the presentation of such a shift space generated as follows [2, Theorem 3.1.5].

The first step is to construct the $(M+1)$th higher edge graph of the full shift $\mathcal{A}^{\mathbb{Z}}$, denoted by $G_{\mathcal{A}^{[M+1]}}$, as follows. The vertex set of $G_{\mathcal{A}^{[M+1]}}$ is the set of all M-blocks over $\mathcal{A}$. For any two vertices $I = a_1 a_2 \cdots a_{M-1} a_M$ and $J = b_1 b_2 \cdots b_{M-1} b_M$ in $G_{\mathcal{A}^{[M+1]}}$, there is exactly one edge from I to J if and only if $a_2 \dots a_M = b_1 \cdots b_{M-1}$. This edge is identified by the string $a_1 a_2 \cdots a_{M-1} a_M b_M = a_1 b_1 \cdots b_{M-1} b_M$. Otherwise, there is no edge from I to J. The next step is to eliminate all edges whose identification is equal to one of the forbidden blocks. After finishing this process, an essential subgraph is taken and the resulting graph is $G_{X^{[M+1]}}$. The vertex set of $G_{X^{[M+1]}}$ is $\mathcal{B}_M(X)$. It is possible to conclude from this construction that a walk on $G_{X^{[M+1]}}$ is a bi-infinite sequence of $(M+1)$-blocks in $\mathcal{B}_{M+1}(X)$ which overlap by M symbols, then $G_{X^{[M+1]}}$ presents the shift space $X^{[M+1]}$. We construct the labeled graph $\mathcal{G}_{X^{[M+1]}} = (G_{X^{[M+1]}}, \mathcal{L}_X)$, by defining the labeling $\mathcal{L}_X : X^{[M+1]} \to \mathcal{A}$ that eliminates the first M symbols from the identification of each edge of $G_{X^{[M+1]}}$. The label of an edge $e = a_1a_2 \ldots a_{M-1}a_Mb_M$ is $\mathcal{L}(e) = b_M$. We discuss next the main properties of the labeled graph $\mathcal{G}_{X^{[M+1]}}$ that presents an M-step shift of finite type.

Property 1: For each vertex $I \in \mathcal{V}(G_{X^{[M+1]}})$, all incoming edges to I have the same label and all outgoing edges from I are labeled distinctly. Then, $\mathcal{G}_{X^{[M+1]}}$ is right-resolving.

Proof: The outgoing edges from $I = a_1a_2 \cdots a_{M-1}a_M$ end at vertices $a_2 \cdots a_{M-1}a_Mb_M$, for $b_M \in \mathcal{A}$. Such an edge (labeled b_M) exists if $a_1a_2 \cdots a_Mb_M \in \mathcal{B}(X)$. So, all outgoing edges from I (at most $|\mathcal{A}|$) have different labels and terminate at different vertices. Analogously, all incoming edges to $J = b_1b_2 \cdots b_M$, $J \in \mathcal{V}(G_{X^{[M+1]}})$, have the same label, b_M, and start at different vertices ($a_1b_1b_2 \cdots b_M$, $a_1 \in \mathcal{A}$). There are no parallel edges in $\mathcal{G}_{X^{[M+1]}}$. ■

Property 2: All M-blocks in $\mathcal{B}_M(X)$ are synchronizing blocks for $\mathcal{G}_{X^{[M+1]}}$. If a block $w = b_1b_2 \ldots b_M$, $w \in \mathcal{B}(X)$, focuses to J, $J \in \mathcal{V}(G_{X^{[M+1]}})$, then $J = b_1b_2 \cdots b_M$. Furthermore, two vertices I_1, I_2 in $\mathcal{G}_{X^{[M+1]}}$ are equivalent if $F^M_{\mathcal{G}_{X^{[M+1]}}}(I_1) = F^M_{\mathcal{G}_{X^{[M+1]}}}(I_2)$.

Proof: If it is possible to reach $J = b_1b_2 \cdots b_M$ from some vertex $I = a_1a_2 \cdots a_M$ in M transitions, the path from I to J must be as follows

$$\begin{array}{ccc} I = a_1a_2 \cdots a_{M-1}a_M & \xrightarrow{b_1} & a_2a_3 \cdots a_Mb_1 \\ a_2a_3 \cdots a_Mb_1 & \xrightarrow{b_2} & a_3a_4 \cdots b_1b_2 \\ \vdots & \vdots & \vdots \\ a_Mb_1 \cdots b_{M-2}b_{M-1} & \xrightarrow{b_M} & b_1b_2 \cdots b_{M-1}b_M = J. \end{array}$$

The block w generated by this process is the same as the block of length M which identifies J, for all initial vertices I (an initial vertex is possible if all transitions shown above are allowed by the language). If $F^M_{\mathcal{G}_{X^{[M+1]}}}(I_1) = F^M_{\mathcal{G}_{X^{[M+1]}}}(I_2)$ the vertices I_1 and I_2 have the same follower set because any M-block generated by a path starting at either I_1 or I_2 reaches the same terminal vertex. ■

It is known that all sofic shifts presented by irreducible labeled graphs are irreducible. However, the converse is not true. The next property proves the converse for $\mathcal{G}_{X^{[M+1]}}$.

Property 3: An M-step shift of finite type X is irreducible if and only if $\mathcal{G}_{X^{[M+1]}} = (G_{X^{[M+1]}}, \mathcal{L}_X)$ is irreducible.

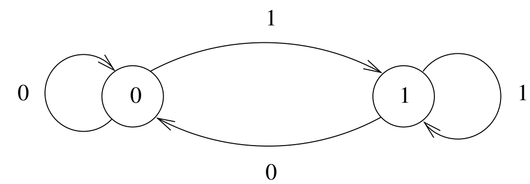

Fig. 1. Out-edge equivalent vertices.

Proof: By Lemma 1, if X is irreducible, then $X^{[M+1]}$ is irreducible. Since $X^{[M+1]}$ is irreducible, the underlying graph $G_{X^{[M+1]}}$ is irreducible by [2, Proposition 2.2.14]. ■

Having established an initial presentation for a shift of finite type, we describe next the minimization algorithm.

A. Graph Minimization

Although there are several labeled graphs which presents the same sofic shift, it is often desirable to find a right-resolving presentation with the smallest number of vertices. The Shannon cover is the unique minimal presentation (up to labeled graph isomorphism) of an irreducible sofic shift that is right-resolving, irreducible, and follower-separated. Given an arbitrary presentation $\mathcal{G} = (G, \mathcal{L})$ of a sofic shift, we can construct another presentation with fewer vertices by merging in $\mathcal{G}$ equivalent vertices. The vertex-minimizing algorithms (see, for example,[2, p. 92], [3, p. 1661]) provide a recursive procedure to determining all classes of equivalent vertices in a right-resolving labeled graph. We will show in this section that the presentation $\mathcal{G}_{X^{[M+1]}}$ allows a very simple recursive identification of the equivalent classes. This identification relies on the concept of *out-edge equivalent* vertices.

Definition 1: Let $\mathcal{G} = (G, \mathcal{V})$ be a labeled graph over an alphabet $\mathcal{A}$. Let $\mathcal{I}$ be a set of all vertices in $\mathcal{V}(G)$ possessing the following property: If there is an edge $I \xrightarrow{a} J$ for some $I \in \mathcal{I}$ and for some $a \in \mathcal{A}$, then there are edges from all other vertices in $\mathcal{I}$ to the same terminal vertex J labeled a. The vertices in $\mathcal{I}$ are called *out-edge equivalent*, and $\mathcal{I}$ is an out-edge equivalence class of $\mathcal{G}$.
Figure 1 shows a labeled graph with one class with two out-edge equivalent vertices. If two vertices are in different out-edge equivalence classes, then they are called out-edge separated vertices, and if all out-edge equivalence classes of a labeled graph $\mathcal{G}$ have only one element, then $\mathcal{G}$ is called out-edge separated graph.

Hereafter, we refer to out-edge equivalent classes with more than one element. A set of out-edge equivalent vertices can be merged by an operation called *in-amalgamation*, defined as follows.

Definition 2: Let $\mathcal{I} = \{I_1, I_2, \cdots, I_m\}$ be an out-edge equivalence class of a labeled graph $\mathcal{G}$. The goal is to create a new labeled graph $\mathcal{H}$ from $\mathcal{G}$ by means of a operation called *in-amalgamation of an out-edge equivalence class* that replaces all vertices within an out-edge equivalent class with one representative of this set, say I_1. This operation redirects into I_1 all edges incoming to $\{I_2, \cdots, I_m\}$, and eliminates the vertices $\{I_2, \cdots, I_m\}$.
The following proposition states that $\mathcal{G}$ and $\mathcal{H}$ presents the same sofic shift.

Proposition 1: Let $\mathcal{G}$ be a labeled graph and $\mathcal{H}$ be a labeled graph obtained from $\mathcal{G}$ by an in-amalgamation of an out-edge equivalence class. Then, $\mathcal{G}$ and $\mathcal{H}$ present the same sofic shift. Furthermore, if $\mathcal{G}$ is irreducible, then so is $\mathcal{H}$, and if $\mathcal{G}$ is right-resolving, then so is $\mathcal{H}$.

Proof: Let $\mathcal{I} = \{I_1, I_2, \cdots, I_m\}$ be an out-edge equivalent class of $\mathcal{G}$. Let $w \in \mathcal{B}(X_\mathcal{G})$ be a block presented by a path π in $\mathcal{G}$. By the construction of $\mathcal{H}$, there is a path π' in $\mathcal{H}$, so that $\mathcal{L}(\pi') = w$, passing through the same vertices of π, except whenever π reaches any vertex $I \in \mathcal{I}$, π' reaches I_1. Then $\mathcal{B}(X_\mathcal{G}) \subseteq \mathcal{B}(X_\mathcal{H})$. Let π be a path in $\mathcal{H}$ presenting the block $v \in \mathcal{B}(X_\mathcal{H})$. By exchanging the roles of I_1 and $I \in \mathcal{I}$ we may use a similar argument to create a path π' in $\mathcal{G}$ presenting v. Then, $\mathcal{B}(X_\mathcal{H}) \subseteq \mathcal{B}(X_\mathcal{G})$.

If $\mathcal{G}$ is irreducible, the previous argument guarantees the existence of a path connecting two arbitrary vertices in $\mathcal{H}$. Clearly, $\mathcal{H}$ inherits the right-resolving property of $\mathcal{G}$ because the labels of the outgoing edges of each vertex in $\mathcal{H}$ are the same as that of the original graph $\mathcal{G}$. ■

Suppose that an arbitrary presentation $\mathcal{G} = (G, \mathcal{L})$ of a sofic shift has t out-edge equivalence classes, namely, $\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_t$. An operation called a *round* of in-amalgamations of out-edge equivalence classes of $\mathcal{G}$ produces a new graph, say $\mathcal{H}_1$, by successively applying the procedure described in Definition 2 t times (one application for each class). We may repeat this process p times until we end up with an out-edge separated graph $\mathcal{H}_p$. The main result of this section is that $\mathcal{H}_p$ is follower-separated if the initial presentation satisfies the conditions stated in the next theorem.

Before we proceed, let us introduce a concept that will be useful in the proof of the next theorem. If all paths that generate a block w, starting from vertices in a set, say $\mathcal{K}$, terminate at the same vertex, this block is called a synchronizing block for $\mathcal{K}$. The set of all blocks of length n generated by the set of vertices $\mathcal{K}$ is denoted by $F^n_\mathcal{G}(\mathcal{K})$.

Theorem 1: Let $\mathcal{G} = (G, \mathcal{L})$ be a right-resolving presentation of a sofic shift with r follower-set equivalence classes, namely $\mathcal{J}_1 \cdots \mathcal{J}_r$, where $\bigcup_i \mathcal{J}_i = \mathcal{V}(G)$. If there exist nonnegative integers $\ell_1, \ell_2, \cdots, \ell_r$ such that all blocks of length ℓ_i in $F^{\ell_i}_\mathcal{G}(\mathcal{J}_i)$ are synchronizing blocks for the ith class, $i = 1, \cdots, r$, the out-edge separated graph $\mathcal{H}_p$ obtained from $\mathcal{G}$ is follower-separated.

Proof: Suppose that there exists an integer ℓ_i such that all blocks in $F^{\ell_i}_\mathcal{G}(\mathcal{J}_i)$ are synchronizing blocks for the class $\mathcal{J}_i$. Let $w \in F^{\ell_i - 1}_\mathcal{G}(\mathcal{J}_i)$ be a block of length $\ell_i - 1$ and $T(w)$ be the set of terminal vertices of all paths presenting w starting at vertices in $\mathcal{J}_i$. Each synchronizing block wa, for $a \in \mathcal{A}$, must be presented by a path starting at each vertex in $\mathcal{J}_i$, because these vertices are equivalent. Then, the vertices in $T(w)$ are out-edge equivalent and they can be merged by an in-amalgamation. After performing these merging operations for all sets $T(w)$, that is, for all $w \in F^{\ell_i - 1}_\mathcal{G}(\mathcal{J}_i)$, all blocks in $F^{\ell_i - 1}_\mathcal{G}(\mathcal{J}_i)$ become synchronizing blocks for the class $\mathcal{J}_i$. We can apply this procedure iteratively until all blocks in $F^1_\mathcal{G}(\mathcal{J}_i)$ become synchronizing blocks for $\mathcal{J}_i$, that is, $\mathcal{J}_i$ becomes an out-edge equivalence class, for $i = 1, \cdots, r$. The last round of in-amalgamations yields the labeled graph $\mathcal{H}_p$ with r states by merging the states within each class $\mathcal{J}_i$, $i = 1, \cdots, r$. Therefore, the equivalent vertices in $\mathcal{V}(G)$ will be merged by successive rounds of in-amalgamations. ■

TABELA I

CONSTRUCTION OF A PRESENTATION FOR A SHIFT OF FINITE TYPE WITH MEMORY M.

1. Find $\mathcal{G}_{X^{[M+1]}}$ from the set $\mathcal{F}$. Do $n = 0$ and $\mathcal{H}_0 = \mathcal{G}_{X^{[M+1]}}$.
2. Identify the out-edge equivalence classes of $\mathcal{H}_n$, $\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_{t_n}$ according to Definition 1.
3. If $\mathcal{H}_n$ is out-edge separated
   then $\mathcal{H}_n$ is the Shannon cover (this graph is unique if the shift is irreducible).
   else Apply a round of in-amalgamation (for those classes with more than one vertex).
   Do $n = n + 1$ and go to step 2.

So, the successive rounds of in-amalgamations is an efficient recursive procedure to identify the equivalence classes of the graph $\mathcal{G}_{X^{[M+1]}}$. Table I summarizes an algorithm to construct a graph presenting a shift of finite type.

Example 1: Let X be a 4-step shift of finite type over $\mathcal{A} = \{0, 1\}$, specified by the set $\mathcal{F} = \{000,\ 1010,\ 1011\}$. Alternatively, $\mathcal{F} = \{0000,\ 0001,\ 1000,\ 1010,\ 1011\}$.
The 4th higher edge graph of the full shift $G_{\mathcal{A}^{[4]}}$ is shown in Fig. 2. After eliminating five edges of $G_{\mathcal{A}^{[4]}}$ and taking the essential subgraph, yields the labeled graph $\mathcal{G}_{X^{[4]}}$ illustrated in Fig. 3. For convenience, the vertex set of $\mathcal{G}_{X^{[4]}}$ is enumerated from 0 to $|\mathcal{V}(\mathcal{G}_{X^{[4]}})| - 1$.

First round: Merge two out-edge equivalence classes of $\mathcal{H}_0$, $\mathcal{I}_1 = \{1, 4\}$, $\mathcal{I}_2 = \{2, 5\}$. The resulting labeled graph $\mathcal{H}_1$ is shown in Fig. 4.

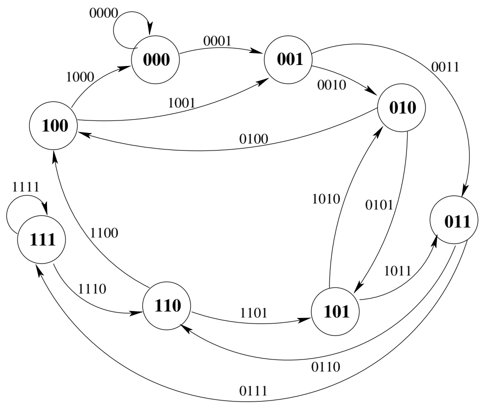

Fig. 2. Forth higher edge graph of the full shift, $G_{\mathcal{A}^{[4]}}$.

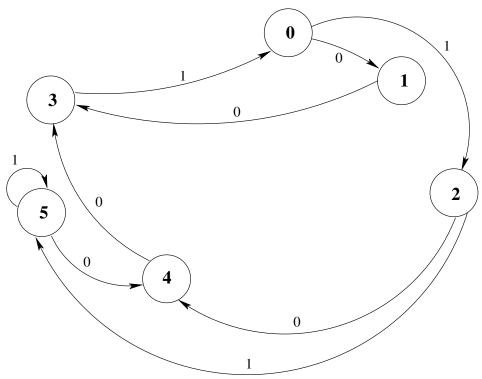

Fig. 3. Initial labeled graph $\mathcal{H}_0 = \mathcal{G}_{X^{[4]}}$.

Second round: Merge one out-edge equivalence class of $\mathcal{H}_1$, $\mathcal{I}_1 = \{0, 2\}$. The labeled graph $\mathcal{H}_2$ shown in Fig. 5 is out-edge separated. So, the minimization procedure has finished and $\mathcal{H}_2$ is the Shannon cover.

IV. Conclusions

In this paper, relevant background material from Symbolic Dynamics has been summarized. The problem of finding the Shannon cover, i.e., a right-resolving presentation with the fewest vertices, from an initial presentation for a shift of finite type has been described. An efficient iterative vertex-minimization algorithm that considers the higher edge graph as the initial graph has been demonstrated. The main practical implication is that the Shannon cover of any shift of finite type can be obtained in an efficient way, benefitting many areas such as coding for storage systems.

References

[1] B. H. Marcus and P. H. Siegel and J. K. Wolf, "Finite-state modulation codes for data storage," *IEEE Journal on Selected Areas in Communications*, vol.10, pp. 5-37, January 1992.

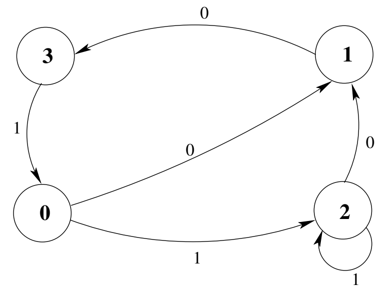

Fig. 4. The labeled graph $\mathcal{H}_1$.

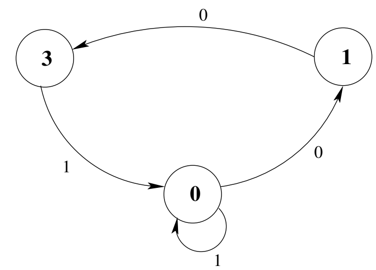

Fig. 5. The labeled graph $\mathcal{H}_2$ (the Shannon cover).

[2] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*. New Jersey, Cambridge University Press, 1995.

[3] B. Marcus, R. Roth, and P. Siegel, "Constrained systems and coding for recording channels," in *Handbook of Coding Theory*, V. S. Pless and W. Huffman, Eds. Amsterdam:Elsevier, 1999, pp. 1635-1764.