Introduction to Symbolic Dynamics

Part 5: The finite-state coding theorem

Silvio Capobianco

Institute of Cybernetics at TUT

May 19, 2010

Revised: November 17, 2010

Overview

- Cyclic structure of irreducible matrices
- Road-colorings and right-closures
- The finite-state coding theorem

Entropy

Definition

The entropy of a nonempty shift X is

$$h(X) = \lim_{n \to \infty} \frac{1}{n} \log |\mathcal{B}_n(X)| = \inf_{n \geq 1} \frac{1}{n} \log |\mathcal{B}_n(X)|$$

If $X = \emptyset$ we put $h(X) = -\infty$.

Basic facts on entropy

- If Y is a factor of X then $h(Y) \leq h(X)$.
- If Y embeds into X then $h(Y) \leq h(X)$.
- If $\mathcal{G} = (G, \mathcal{L})$ is right-resolving then $h(\mathrm{X}_{\mathcal{G}}) = h(\mathrm{X}_G)$.

The Perron-Frobenius theorem

Let A be a nonnegative irreducible nonzero matrix.

1. A has a positive eigenvector $\mathbf{v}_A$.
2. The eigenvalue λ_A corresponding to $\mathbf{v}_A$ is positive.
3. λ_A is algebraically—and geometrically—simple, *i.e.*,
   - $\det(tI - A) = (t - \lambda_A)p(t)$ with $p(\lambda_A) \neq 0$, and
   - $\dim\{\mathbf{v} \mid A\mathbf{v} = \lambda_A\mathbf{v}\} = 1$.
4. If μ is another eigenvalue of A then $|\mu| \leq \lambda_A$.
5. Any positive eigenvector of A is a positive multiple of $\mathbf{v}_A$.

The value λ_A is called the Perron eigenvalue of A

Computing entropy via the Perron-Frobenius theorem

Theorem

- Let G be a graph, let A be its adjacency matrix, and let λ_A be the maximum Perron eigenvalue of an irreducible component of A.
- Then $h(\mathsf{X}_G) = \log \lambda_A$.
- In addition, if $\mathcal{G} = (G, \mathcal{L})$ is right-resolving, then $h(\mathsf{X}_{\mathcal{G}}) = \log \lambda_A$.

Periods

Period of a shift

If X is a shift we define

$$\mathrm{per}\, X = \gcd\{n \in \mathbb{N} \mid p_n(X) > 0\}$$

with the conventions $\gcd \emptyset = \infty$, $\gcd(U \cup \{\infty\}) = \gcd U$.

Period of a matrix

Let G be graph and A its adjacency matrix. The period of a state I is

$$\mathrm{per}\, I = \gcd\{n \in \mathbb{N} \mid (A^n)_{I,I} > 0\}$$

The period of A (and G) is

$$\mathrm{per}\, G = \mathrm{per}\, A = \gcd\{\mathrm{per}\, I \mid I \in \mathcal{V}(G)\} = \mathrm{per}\, \mathsf{X}_G$$

A is aperiodic if $\mathrm{per}\, A = 1$.

Periods of irreducible graphs

Theorem

States of an irreducible graph have same period.

Reason why

- Suppose $p = \mathrm{per}\, I$ and n is a period of J.
- Suppose $(A^r)_{I,J} > 0$ and $A^s_{J,I} > 0$.
- Then p divides both $r + s$ and $r + n + s$...

Period equivalence

Definition

- Let G be an irreducible graph s.t. $A = A(G)$ is nonzero.
- States I and J are period equivalent if there is a path from I to J whose length is divisible by $\operatorname{per} G$.

Period equivalence is an equivalence relation

A path from I to J plus a path from J to I form a cycle from I to I.

Period classes

A period class is a class of period equivalence.

Periodic decomposition

Theorem

Let A be an irreducible nonzero matrix and let p be its period.

- Period equivalence on A has p classes.
- There is an ordering $D_0, \ldots, D_{p-1}$ of period classes s.t. every edge e with $\mathrm{i}(e) \in D_i$ has $\mathrm{t}(e) \in D_{(i+1) \bmod p}$.

Proof

- Fix D_0 and just put $D_{i+1} = \{\mathrm{t}(e) \mid \mathrm{i}(e) \in D_i\}$.
- By construction, each D_i is a period class. There are p of them because A is irreducible. Each edge from D_{p-1} must end in D_0.

Cyclic form of an irreducible nonzero matrix

By previous argument, after renaming the states,

$$A = \begin{pmatrix} 0 & B_0 & 0 & \dots & 0 \\ 0 & 0 & B_1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & B_{p-2} \\ B_{p-1} & 0 & 0 & \dots & 0 \end{pmatrix}$$

Moreover,

$$A^p == \begin{pmatrix} A_0 & 0 & 0 & \dots & 0 \\ 0 & A_1 & 0 & \dots & 0 \\ 0 & 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & A_p \end{pmatrix}$$

for suitable A_i's.

Primitive graphs

Definition

- A matrix is primitive if it is irreducible and aperiodic.
- A graph is primitive if its adjacency matrix is primitive.

Characterization

Let A be a nonnegative matrix. TFAE.

1. A is primitive.
2. A^N is positive for some N.
3. A^N is positive for all sufficiently large N.

Rationale

- If A is primitive, then $(A^n)_{I,I} > 0$ for all $n \geq N_I$.
- Put $N = M + \max_{i \in \mathcal{V}} N_I$ where $(A^n)_{I,J} > 0$ for some $n \leq M$.

Mixing shifts

Definition

A shift X is mixing if for any $u, v \in \mathcal{B}(X)$ there exists $N \geq 1$ s.t. for every $n \geq N$ there exists $w \in \mathcal{B}_n(X)$ s.t. $uwv \in \mathcal{B}(X)$.

Facts

- A factor of a mixing shift is mixing.
- If G is essential then X_G is mixing iff G is primitive.
- A SFT is mixing iff it is irreducible and aperiodic.
- For a mixing sofic shift,

$$\lim_{n \to \infty} \frac{1}{n} \log p_n(X) = \lim_{n \to \infty} \frac{1}{n} \log q_n(X) = h(X)$$

Road-colorings

Definition

- Let $G = (\mathcal{V}, \mathcal{E})$ a graph. Recall that $\mathcal{E}_I = \{e \in \mathcal{E} \mid \mathrm{i}(e) = I\}$.
- A labeling $\mathcal{C} : \mathcal{E} \to A$ is a road-coloring if it is bijective on each $\mathcal{E}_I$.
- A graph G is road-colorable if it admits a road-coloring.

Characterization

Road-colorable graphs are precisely those with constant out-degree.

Use

- Observe that a road-coloring is right-resolving.
- Given a word w over A and a state I in G, there is exactly one path from I labeled w.
- In particular, $(G, \mathcal{C})$ is a presentation of the full A-shift.

The road-coloring problem

Statement

Is it true that every road-colorable primitive graph has a road-coloring admitting a synchronizing word?

Status at time of publication of Lind and Marcus textbook

Unsolved.

Current status

Solved.

- Trahtman, Avraham N. (2009) The road colouring problem. *Israel Journal of Mathematics* **172(1)**: 51–60.

Thanks to Prof. Trahtman for correction. (2010-11-17)

Right-closing graphs

Definition

- Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph.
- Suppose that, given any two paths $\pi = \pi_1 \dots \pi_{D+1}$ and $\rho = \rho_1 \dots \rho_{D+1}$ of length $D+1$, if $\mathrm{i}(\pi) = \mathrm{i}(\rho)$ and $\mathcal{L}(\pi) = \mathcal{L}(\rho)$, then $\pi_1 = \rho_1$.
- We then say that $\mathcal{G}$ is right-closing with delay D.

Motivation

- $\mathcal{G}$ is right-resolving iff it is right-closing with delay zero.
- Two paths of length $N > D$ on a right-closing graph, that have same labeling and same initial state, are equal for the first $N - D$ steps.

One-sided shifts

Definition

If X is a (two-sided) shift over A, we put

$$X^+ = \{x_{[0,\infty)} \mid x \in X\}$$

Special cases

- If $X = \mathsf{X}_G$, then X^+ is the set of infinite paths on G.
- If $X = \mathsf{X}_{\mathcal{G}}$, then X^+ is the set of labelings of infinite paths on $\mathcal{G}$.
- The map $\mathcal{L}_\infty^+ : \mathsf{X}_G^+ \to \mathsf{X}_{\mathcal{G}}^+$ defined by $\mathcal{L}^+(\pi)_i = \mathcal{L}(\pi_i)$ is surjective.

Characterization of right-closing graphs

Theorem

Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph and let $\mathrm{X}^+_{G,I} = \{\pi \in \mathrm{X}^+_G \mid \mathrm{i}(\pi) = I\}$.
TFAE.

1. $\mathcal{G}$ is right-closing.
2. For every state I, $\mathcal{L}^+ : \mathrm{X}^+_{G,I} \to \mathrm{X}^+_{\mathcal{G}}$ is injective.

Reason why

- Suppose $\mathcal{G}$ is not right-closing.
- For $n > |\mathcal{V}|^2$ find π and ρ of same length n, same initial state, and different initial edge.
- Then $\pi = \alpha_1\alpha_2\alpha_3$, $\rho = \beta_1\beta_2\beta_3$ with $|\alpha_i| = |\beta_i|$ and α_2 and β_2 loops.
- Then $\mathcal{L}^+(\alpha_1(\alpha_2)^\infty) = \mathcal{L}^+(\beta_1(\beta_2)^\infty)$.

Conditions on right-closure

A sufficient condition

- Let $\mathcal{G} = (G, \mathcal{L})$ be s.t. $\mathcal{L}_\infty$ is a conjugacy.
- Suppose $\mathcal{L}_\infty^{-1}$ has anticipation n.
- Then $\mathcal{L}$ is right-closing with delay n.

A necessary condition

- Let $\mathcal{G} = (G, \mathcal{L})$ be right-closing with delay D.
- Let $\mathcal{H}$ be obtained from $\mathcal{G}$ via out-splitting.
- Then $\mathcal{H}$ is right-closing with delay $D + 1$.

Reasons why

- We can always suppose G essential, so every path is left-extendable.
- Splitting has memory 0 and anticipation 1; amalgamation is 1-block.

Right-closing labelings preserve entropy

Theorem

- Let $\mathcal{G} = (G, \mathcal{L})$ be a labeled graph.
- Suppose $\mathcal{L}$ is right-closing.
- Then $h(\mathsf{X}_{\mathcal{G}}) = h(\mathsf{X}_G)$.

Reason why

- Initial state and labeling of a $D+1$-path determine first edge.
- Thus, if G has r states, then $|\mathcal{B}_n(\mathsf{X}_G)| \leq r \cdot |\mathcal{B}_{n+D}(\mathsf{X}_{\mathcal{G}})|$.

Recoding right-closure into right-resolvedness

Theorem

Let $\mathcal{G} = (G, \mathcal{L})$ be a right-closed labeled graph with delay D. There exist a graph H and labelings Ψ, Θ on H s.t.

$$\begin{tikzcd}
\mathsf{X}_G \arrow[d, "\mathcal{L}_\infty"'] & & \mathsf{X}_H \arrow[ll, "\Theta_\infty \circ \sigma^D"'] \arrow[dll, "\Psi_\infty"] \\
\mathcal{L}_\infty(\mathsf{X}_G) & &
\end{tikzcd}$$

with Ψ right-resolving and Θ a conjugacy.

Reason why (for $D > 0$)

- Put $\mathcal{V}(H) = \{(I, \mathcal{L}(\pi)) \mid I \in \mathcal{V}(G), \mathrm{i}(\pi) = I, |\pi| = D\}$.
- An edge in H joins $(I, \mathcal{L}(\pi))$ to $(\mathrm{t}(e), \mathcal{L}(\pi_{[2,D]})a)$ where I and $\mathcal{L}(\pi)a$ determine $e \in \mathcal{E}(G)$. Call $(I, \mathcal{L}(\pi)a)$ such edge.
- Put $\Theta(I, \mathcal{L}(\pi)a) = e$. Put $\Psi(I, \mathcal{L}(\pi)a) = a$.

Finite-state codes

Definition

A finite-state code is a triple $(G, \mathcal{I}, \mathcal{O})$ where:

- G is a graph—encoder graph
- $\mathcal{I}$ is a road-coloring on G—input labeling
- $\mathcal{O}$ is a right-closing labeling on G—output labeling

A finite-state (X, n)-code is a finite-state code where:

- G has out-degree n.
- $\mathcal{O}_\infty(\mathsf{X}_G) \subseteq X$.

Using finite-state codes

Drawing finite-state codes as labeled graphs

Edge e is marked as $\mathcal{I}(e)/\mathcal{O}(e)$. Example:

$$\begin{tikzcd} \bullet \arrow[loop left, "1/c"] \arrow[r, bend left, "0/a"] & \bullet \arrow[l, bend left, "1/c"] \arrow[loop right, "0/b"] \end{tikzcd}$$

Encoding sequences on n-ary alphabets

- Let $(G, \mathcal{I}, \mathcal{O})$ be a finite-state (X, n)-code
- Let $x_0 x_1 x_2 \ldots$ be an infinite sequence on an n-ary alphabet.
- Fix $I_0 \in \mathcal{V}(G)$. There is exactly one sequence $e_0 e_1 e_2 \ldots$ of edges s.t. $\mathcal{I}(e_i) = x_i$ for every i.
- The same sequence is also encoded as $\mathcal{O}(e_0)\mathcal{O}(e_1)\mathcal{O}(e_2)\ldots \in X^+$.
- Since $\mathcal{O}$ is right-closing, input can be reconstructed from output, given the initial state.

The finite-state coding theorem

Statement

Let X be a sofic shift. TFAE.

1. There exists a finite-state (X, n)-code.
2. $h(X) \geq \log n$.

Necessity of the condition

- $h(\mathsf{X}_G) = h(\mathcal{I}_\infty(\mathsf{X}_G)) = h(\mathcal{O}_\infty(\mathsf{X}_G))$ because $\mathcal{I}$ and $\mathcal{O}$ are right-closing.
- $h(\mathcal{I}_\infty(\mathsf{X}_G)) = \log n$ because $(G, \mathcal{I})$ is a presentation of the full n-shift.
- $h(\mathcal{O}_\infty(\mathsf{X}_G)) \leq h(X)$ because $\mathcal{O}_\infty(\mathsf{X}_G) \subseteq X$.

Enforcing finite-state coding

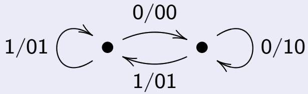
Encoding the full 2-shift into a binary sofic shift

Not possible right away, but. . .

- Divide input into blocks of length p, *i.e.*, use $\mathsf{X}_{[2^p]}$ instead of $\mathsf{X}_{[2]}$.
- Divide output into blocks of length q, *i.e.*, use X^q instead of X.
- Then condition becomes $h(X) \geq p/q$.

Example with the $(1,3)$ run-length limited shift

- $h(\mathsf{X}(1,3)) \approx 0.55$, so we take $p = 1$ and $q = 2$.
- The input alphabet is still the full 2-shift.
- The output alphabet is $\mathcal{B}_2(\mathsf{X}(1,3)) = \{00, 01, 10\}$.
- The labeled graph below yields the modified frequency modulation:



Approximate eigenvectors

Definition

- Let A be a nonnegative, integral matrix.
- Let n be a positive integer.
- Let $\mathbf{v}$ be a nonnegative, nonzero, integral vector.
- $\mathbf{v}$ is an (A, n)-approximate eigenvector if $A\mathbf{v} \geq n\mathbf{v}$.

Example

- Let $A = \begin{pmatrix} 1 & 3 \\ 6 & 1 \end{pmatrix}$.
- Then $\mathbf{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ is an $(A, 5)$-approximate eigenvector.

Interpretations

Physical

- Suppose we assign weight v_I to state I.
- Then $\sum_{\mathrm{i}(e)=I} v_{\mathrm{t}(e)} \geq n \cdot v_I$ for every state I.

Geometrical

- Suppose A is an $r \times r$ matrix.
- Each inequality $\sum_{J=1}^{r} A_{I,J} x_J \geq n \cdot x_I$ determines a closed half-space.
- Then, (A, n)-approximate eigenvectors are elements of a closed cone in r-dimensional space.

Positive approximate eigenvectors

Lemma

- Let G be a graph and $A = A(G)$ its adjacency matrix.
- Let $\mathbf{v}$ be an (A, n)-approximate eigenvector.
- Then there exists a subgraph H of G s.t.

$$\mathbf{w}_I = \mathbf{v}_i \ \ \forall I \in \mathcal{V}(H)$$

is a positive $(A(H), n)$-approximate eigenvector.

Reason why

- Let K be the subgraph generated by the states where $v_I > 0$.
- K has an irreducible component H which is a sink.

Looking for approximate eigenvectors

Theorem

Let A be a nonnegative matrix. TFAE.

1. There exists an (A, n)-approximate eigenvector.
2. $\lambda_A \geq n$.

Moreover, if A is irreducible then there exists a positive (A, n)-approximate eigenvector.

Reason why

- It is not restrictive that A is irreducible and $\mathbf{v}$ positive.
- If $\mathbf{v}$ is an (A, n)-approximate eigenvector then $c, d > 0$ exist s.t. $cn^k \leq \sum_{I,J=1}^{r} (A^k)_{I,J} \leq d\lambda_A^k$ for every k, thus $n \leq \lambda_A$.
- If $\lambda_A = n$ then $\mathbf{v}_A$ is rational: use a suitable multiple.
- If $\lambda_A > n$ modify $\mathbf{v}_A$ into a rational $\mathbf{v}$ s.t. $A\mathbf{v} > n\mathbf{v}$ still holds.

Finding approximate eigenvectors

Algorithm

INPUT: nonnegative integral A and $\mathbf{z}$, positive integer n.

1. Compute $\mathbf{z}' = \min \left\{ \mathbf{z}, \left\lfloor \frac{1}{n} A\mathbf{z} \right\rfloor \right\}$
2. If $\mathbf{z}' = \mathbf{z}$: return $\mathbf{z}$
3. Replace $\mathbf{z}$ with $\mathbf{z}'$
4. Repeat

OUTPUT: either an (A, n)-approximate eigenvector, or the null vector.

Use

- Put $(\mathbf{v}_k)_I = k$ for every I.
- Apply the algorithm to $\mathbf{v}_1$, then to $\mathbf{v}_2$, and so on, until output is non-null.
- Then the final output is the smallest (A, n)-approximate eigenvector.

Approximate eigenvectors and splittings

Lemma A

- Let G be an irreducible graph and let $A = A(G)$.
- Suppose $\lambda_A \geq n$.
- Then there exists a sequence of graphs

$$G = G_0, G_1, \ldots, G_m = H$$

  such that:
  - Each G_i is an elementary splitting of G_{i-1}.
  - $|\mathcal{E}_I(s)| \geq n$ for every state s in H.
- Let $\mathbf{v}$ be a positive (A, n)-approximate eigenvector, and let $k = \sum_{I \in \mathcal{V}(G)} v_i$.
- Then the sequence above can be chosen with $m \leq k - |\mathcal{V}(G)|$ and $|\mathcal{V}(H)| \leq k$.

Proof of the finite-state coding theorem

- Let $X = \mathsf{X}_{\mathcal{K}}$ be a sofic shift s.t. $h(X) \geq \log n$.
- We may suppose $\mathcal{K} = (K, \mathcal{L})$ irreducible and right-resolving
- If $A = A(K)$ then $\lambda_A = h(X) \geq \log n$.
- Construct a sequence $K = G_0, G_1, \ldots, G_m = H$ s.t.
  - Each G_i is an elementary splitting of G_{i-1}.
  - $|\mathcal{E}_I(s)| \geq n$ for every state s in H.
- The labeling $\mathcal{L}'$ of H resulting from $\mathcal{L}$ is right-closing with delay $\leq m$.
- Construct $(G, \mathcal{I}, \mathcal{O})$ as follows:
  - G is a subgraph of H with constant out-degree n.
  - $\mathcal{I}$ is any road-coloring of G.
  - $\mathcal{O}$ is the restriction of $\mathcal{L}'$ to G.
- Then $(G, \mathcal{I}, \mathcal{O})$ is a finite-state (X, n)-code.

The state splitting algorithm

INPUT: a sofic shift X.

1. Construct a right-resolving presentation $\mathcal{K} = (K, \mathcal{L})$ of X.
2. Compute $h(X) = \log \lambda_{A(K)}$.
3. Choose integers p and q s.t. $h(X) \geq p/q$.
4. Construct $\mathcal{K}^q$—which is a right-resolving presentation of X^q.
5. Use the approximate eigenvector algorithm to find an $(A(K^q), 2^p)$-approximate eigenvector. Then reduce to a sink component $\mathcal{H}$ with positive approximate eigenvector.
6. Perform a chain of state splits until obtaining a presentation with minimum out-degree $\geq 2^p$.
7. Prune to obtain $\mathcal{G} = (G, \mathcal{O})$ with constant out-degree 2^p. Choose a road-coloring $\mathcal{I}$ using binary p-blocks.

OUTPUT: A rate $p : q$ finite-state code $(G, \mathcal{I}, \mathcal{O})$.

Propagation of errors with finite-state codes

Example

- Consider the finite-state code $$\begin{tikzcd} \bullet \arrow[loop left, "0/a"] \arrow[r, bend left, "1/b"] & \bullet \arrow[l, bend left, "0/c"] \arrow[loop right, "1/a"] \end{tikzcd}$$
- If the initial state is the one on the left, $00000\ldots$ is encoded into $aaaaa\ldots$
- However, suppose that an error occurs, and the first a is written b.
- Then a decoder would reconstruct $11111\ldots$

Sliding block decoders

Definition

- Let $(G, \mathcal{I}, \mathcal{O})$ be a finite-state (X, n)-code.
- A sliding block decoder for $(G, \mathcal{I}, \mathcal{O})$ is a SBC $\phi : X \to \mathsf{X}_{[n]}$ s.t.

$$\begin{tikzcd} & \mathsf{X}_G \arrow[dl, "\mathcal{I}_\infty"'] \arrow[dr, "\mathcal{O}_\infty"] & \\ \mathsf{X}_{[n]} & & X \arrow[ll, "\phi"] \end{tikzcd}$$

Use

- Suppose $\phi = \Phi_\infty^{[-m,\alpha]}$. Let $y_0 y_1 y_2 \dots$ be an output sequence.
- For $k \geq m$ it is $y_{k-m} \dots y_{k+\alpha} = \mathcal{O}(e_{k-m} \dots e_{k+\alpha})$.
- Then $x_k = \mathcal{I}(e_k) = \Phi(y_{k-m} \dots y_{k+\alpha})$,

i.e., input can be reconstructed from output without recording the state, except at most the first m symbols.

The sliding block decoding theorem

Statement

- Let X be a shift of finite type.
- Suppose $h(X) \geq \log n$.
- Then there exists an (X, n)-finite state code with a sliding block decoder.

Reason why

The labeling of a minimal right-resolving presentation is a conjugacy.

Consequence

- Let X be a SFT.
- Suppose $h(X) \geq \log n$.
- Then X factors onto the full n-shift.

Thank you for attention!