Introduction to Symbolic Dynamics

Part 1: The basics

Silvio Capobianco

Institute of Cybernetics at TUT

April 14, 2010
Overview

- Historical introduction
- Shift subspaces
- Basic constructions on shift subspaces
- Sliding block codes
- A parallel with coding theory
A short history of symbolic dynamics

- 1898:
  Hadamard’s work on geodetic flows.

- 1930s:
  Morse and Hedlund’s work.

- 1960s:
  Smale introduces the word “subshift”.

- 1990s:
  Boyle and Handelman make a crucial step towards characterization of nonzero eigenvalues of nonnegative matrices.
Hadamard’s problem

Geodesic flows on surfaces of negative curvature
Generally hard problem, but...

What if...
- Partition the space into finitely many regions.
- Discretize time.
- Check the region instead of the exact position.

Discovery!
The complicated dynamics can be described in terms of finitely many forbidden pairs of symbols!
Sequences and blocks

Full shifts

Let $A$ be a finite alphabet. The full $A$-shift is the set

$$A^\mathbb{Z} = \{\text{bi-infinite words on } A\}.$$  

The full $r$-shift is the full $A$-shift for $A = \{0, \ldots, r - 1\}$.

Blocks

- A block, or word, over $A$ is a finite sequence $u$ of elements of $A$.
- If $u = a_1 \ldots a_k$ then $k = |u|$ is the length of $u$. If $|w| = 0$ then $w = \varepsilon$.
- A subblock of $u = a_1 \ldots a_k$ has the form $v = a_i \ldots a_j$, $1 \leq i, j \leq k$.
  If $x \in A^\mathbb{Z}$ then $x_{[i,j]}$ is the subblock $x_i \ldots x_j$.
- A block $u$ occurs in a sequence $x$ if $x_{[i,j]} = u$ for some $i, j \in \mathbb{Z}$. 
The shift map

\[ \sigma(x)_i = x_{i+1} \text{ for all } x \in A^\mathbb{Z}, \ i \in \mathbb{Z}. \]

**Periodic points**
- \( x \in A^\mathbb{Z} \) is **periodic** if \( \sigma^n(x) = x \) for some \( n > 0 \).
- Any such \( n \) is called a **period** of \( x \).
- \( x \) is a **fixed point** for \( \sigma \) if \( \sigma(x) = x \).

**Consequences**
- Definition above is the same as \( x_{i+n} = x_i \ \forall i \in \mathbb{Z} \).
- If \( x \) has a period, then it also has a least period.
Interpretation

- The group $\mathbb{Z}$ represents time.
- (Bi-infinite) sequences represent (reversible) trajectories.
- The shift represents the passing of time.
- Periodic sequences represent periodic (closed) trajectories.
Shift subspaces

**Definition**

Let $\mathcal{F}$ be a set of blocks over $A$ and let

$$X_\mathcal{F} = \left\{ x \in A^\mathbb{Z} \mid x[i,j] \neq u \quad \forall i,j \in \mathbb{Z} \forall u \in \mathcal{F} \right\}$$

A **shift subspace**, or **subshift**, over $A$ is a subset of $A^\mathbb{Z}$ of the form $X = X_\mathcal{F}$ for some set of blocks $\mathcal{F}$. 
Examples of subshifts

1. The full shift.

2. The golden mean shift $X = X_{\{11\}}$.

3. The even shift $X = X_F$ with $F = \{10^{2k+1}1 \mid k \in \mathbb{N}\}$.

4. For $S \subseteq \mathbb{N}$, the $S$-gap shift $X(S)$ with $F = \{10^n1 \mid n \in \mathbb{N} \setminus S\}$. For $S = \{d, \ldots, k\}$ we have the $(d,k)$ run-length limited shift $X(d,k)$.

5. The set of labelings of bi-infinite paths on the graph

\[
\begin{array}{c}
\bullet \\
\begin{array}{c}
\text{e} \quad \text{f} \\
\text{g}
\end{array}
\end{array}
\]

6. The charge constrained shift over $\{+1, -1\}$ s.t. $x \in X$ iff $\sum_{i=j}^{j+n} x_i \in [-c, c]$ for every $j \in \mathbb{Z}$, $n \geq 0$.

7. The context free shift over $\{a, b, c\}$ with

$$F = \{ab^m c^k a \mid m \neq k\}$$
Basic facts on subshifts

1. Suppose $X_1 = X_{\mathcal{F}_1}$ and $X_2 = X_{\mathcal{F}_2}$. Then $X_1 \cap X_2 = X_{\mathcal{F}_1 \cup \mathcal{F}_2}$.

2. Suppose $\mathcal{F}_1 \subseteq \mathcal{F}_2$. Then $X_{\mathcal{F}_1} \supseteq X_{\mathcal{F}_2}$. In particular, $X_1 \cup X_2 \subseteq X_{\mathcal{F}_1 \cap \mathcal{F}_2}$.

3. In general, $X_1 \cup X_2 \neq X_{\mathcal{F}_1 \cap \mathcal{F}_2}$.

4. Let $\{X_i\}_{i \in I}$ be a family of subshifts s.t. $\bigcup_{i \in I} X_i = A^\mathbb{Z}$. Then $X_i = A^\mathbb{Z}$ for some $i \in I$.

5. If $X$ is a subshift over $A$ and $Y$ is a subshift over $B$, then

$$X \times Y = \{z : \mathbb{Z} \to A \times B \mid \exists x \in X, y \in Y \mid \forall i \in \mathbb{Z}. z_i = (x_i, y_i)\}$$

is a subshift over $A \times B$. 
Shift invariance

**Definition**

$X \subseteq A^\mathbb{Z}$ is **shift invariant** if $\sigma(X) \subseteq X$.

**Subshifts are shift invariant**

Write $\sigma_X$ for the restriction of the shift to $X$.

**Shift invariance is not enough to make a subshift!**

$$X = \left\{ x \in \{0, 1\}^\mathbb{Z} \mid \exists! i \mid x_i = 1 \right\}$$

- $X$ is shift invariant.
- And no block of the form $0^n$ is forbidden.
- Then, if $X$ were a subshift, it would contain $0^\mathbb{Z}$—which it doesn’t.
Languages

Definition

Let $X \subseteq A^\mathbb{Z}$, not necessarily a subshift.
Let $B_n(X)$ be the set of subblocks of length $n$ of elements of $X$. The language of $X$ is

$$B(X) = \bigcup_{n \geq 0} B_n(X).$$
Characterization of subshift languages

1. Let $X$ be a subshift. Let $L = B(X)$.
   1. For every $w \in L$, if $u$ is a factor of $w$, then $u \in L$.
   2. For every $w \in L$ there exist nonempty $u, v \in L$ s.t. $uwv \in L$.

2. Suppose $L \subseteq A^*$ satisfies points 1 and 2 above.
   Then $L = B(X)$ for some subshift $X$ over $A$.

3. In fact, if $X$ is a subshift and $L = B(X)$, then $X = X_{A^* \setminus L}$.
   In particular, the language of a subshift determines the subshift.

4. Subshifts over $A$ are precisely those $X \subseteq A^\mathbb{Z}$ s.t.
   for every $x \in A^\mathbb{Z}$,
   if $x_{[i,j]} \in B(X)$ for every $i, j \in \mathbb{Z}$,
   then $x \in X$.

5. In particular, a finite union of subshifts is a subshift.
Irreducibility

Definition
A subshift $X$ is **irreducible** if for every $u, v \in B(X)$ there exists $w \in B(X)$ s.t. $uwv \in B(X)$.

Meaning
$X$ is irreducible iff the **dynamical system** $(X, \sigma)$ is not made of two parts not joined by any **orbit**.

Examples
- The golden mean shift is irreducible.
- The subshift $X = \{0^\mathbb{Z}, 1^\mathbb{Z}\}$ is not irreducible.
Higher block shifts

Let $X$ be a subshift over $A$. Consider $A_X^{[N]} = B_N(X)$ as an alphabet.

The $N$-th higher block code

It is the map $\beta_N : X \rightarrow (A_X^{[N]})^\mathbb{Z}$ defined by

$$(\beta_N(x))_i = x_{[i,i+N-1]}$$

The $N$-th higher block shift

It is the subshift $X^{[N]} = \beta_N(X)$. 
Higher block shifts are subshifts

Let $X = X_\mathcal{F}$. It is **not** restrictive to suppose $|u| \geq N$ for every $u \in \mathcal{F}$.

For $|w| \geq N$ put $w_i^{[N]} = w[i:i+N-1]$. Let

$$\mathcal{F}_1 = \{w^{[N]} \mid w \in \mathcal{F}\}.$$

Then put

$$\mathcal{F}_2 = \{uv \mid u, v \in A^N, \exists i > 1 \mid u_i \neq v_{i-1}\}$$

Then clearly $X^{[N]} \subseteq X_{\mathcal{F}_1 \cup \mathcal{F}_2}$. On the other hand, any $x \in X_{\mathcal{F}_1 \cup \mathcal{F}_2}$ reconstructs some $y \in X$, so that $x = \beta_N(y) \in X^{[N]}$. 
Higher power shifts

Let $X$ be a subshift over $A$. Consider $A_X^{[N]} = \mathcal{B}_N(X)$ as an alphabet.

**The $N$-th higher power code**

It is the map $\gamma_N : X \rightarrow (A_X^{[N]})^\mathbb{Z}$ defined by

$$(\gamma_N(x))_i = x_{[Ni,N(i+1)-1]}$$

**The $N$-th higher power shift**

It is the subshift $X_N = \gamma_N(X)$. 
Higher block shifts and other operations

Properties

1. \((X \cap Y)^[N] = X^[N] \cap Y^[N]\).
2. \((X \cup Y)^[N] = X^[N] \cup Y^[N]\).
3. \((X \times Y)^[N] = X^[N] \times Y^[N]\).
4. \(\beta_N \circ \sigma_X = \sigma_{X^N} \circ \beta_N\)

A note on higher power shifts

\(\gamma_N \circ \sigma_X^N = \sigma_{X^N} \circ \gamma_N\).
Sliding block codes

- Let $X$ be a subshift over $A$. Let $\mathcal{A}$ be another alphabet.
- Let $\Phi : B_{m+n+1}(X) \to \mathcal{A}$.
- Then $\phi : X \to \mathcal{A}^\mathbb{Z}$ defined by
  \[
  \phi(x)_i = \Phi\left(x_{[i-m,i+n]}\right)
  \]
  is a sliding block code (SBC) with memory $m$ and anticipation $n$.
- We then write $\phi = \Phi_{[-m,n]}$, or just $\phi = \Phi_{\infty}$.
- We may also write $\phi : X \to Y$ if $Y$ is a subshift over $\mathcal{A}$ and $\phi(X) \subseteq Y$.
- It is always possible to increase both memory and anticipation.
- We speak of 1-block code when $m = n = 0$. 
Examples of sliding block codes

1. The shift.
2. The identity.
3. The converse of the shift.
4. The $N$-th higher block code map $\beta_N$.
5. The xor, induced by $\Phi(x_0x_1) = x_0 + x_1 \mod 2$.
6. The map defined by

   \[ \phi(00) = 1, \phi(01) = 0, \phi(10) = 0 \]

   is a SBC from the golden mean shift to the even shift.
The key property of SBC

Let $X$ and $Y$ be shift spaces, and let $\phi : X \to Y$ be a SBC. Then

$$
\begin{array}{ccc}
X & \xrightarrow{\sigma_X} & X \\
\downarrow \phi & & \downarrow \phi \\
Y & \xrightarrow{\sigma_Y} & Y
\end{array}
$$

Meaning

- SBC are shift-commuting.
- SBC represent stationary processes.
- A SBC from $X$ to $Y$ is a morphism from $(X, \sigma)$ to $(Y, \sigma)$. 
Shift-commutativity is not enough to make a SBC

Counterexample

Let $\phi(x) : \{0, 1\}^\mathbb{Z} \rightarrow \{0, 1\}^\mathbb{Z}$ be defined by

$$
\phi(x)_i = \begin{cases} 
1 - x_i & \text{if } \exists j > i \mid x_j = 1, \\
x_i & \text{otherwise}.
\end{cases}
$$

Theorem

Let $\phi : X \rightarrow Y$ be a map between shift spaces. Then $\phi$ is a SBC if and only if:

1. $\phi$ is shift-commuting, and
2. there exists $N \geq 0$ s.t. $\phi(x)_0$ is a function of $x_{[-N:N]}$.

Consequently, compositions of SBC are SBC.
Factors, embeddings, conjugacies

Let $X$ and $Y$ be subshifts, $\phi : X \to Y$ a SBC.

Factors
- $\phi$ is a factor code if it is surjective.
- $Y$ is a factor of $X$ if there is a surjective SBC from $X$ to $Y$.

Embeddings
- $\phi$ is an embedding if it is injective.

Conjugacies
- $\phi$ is a conjugacy if it is bijective.
- The $N$th higher block code is a conjugacy from $X$ to $X^{[N]}$, with converse
  $$\beta_N^{-1}(y)_i = (y_i)_0.$$
Every SBC can be recoded as a 1-SBC

**Theorem**

For every SBC $\phi : X \rightarrow Y$ there exist an integer $N > 0$, a conjugacy $\psi : X \rightarrow X^{[N]}$, and a 1-block code $\omega : X^{[N]} \rightarrow Y$ such that

\[
\begin{array}{c}
X \\
\downarrow \phi \\
Y
\end{array} \xrightarrow{\psi} \begin{array}{c}
X^{[N]} \\
\downarrow \omega
\end{array}
\]

**Reason why**

- Suppose $\phi = \Phi_{\infty}^{[-m,n]}$.
- Put $N = m + n + 1$, $\psi = \sigma^{-m} \circ \beta_N$.
- Then $\omega = \phi \circ \psi^{-1} = \phi \circ \beta_N^{-1} \circ \sigma^m$ is a 1-SBC.
Theorem

Let $X$ be a shift space over $A$
Let $\phi : X \to A^\mathbb{Z}$ be a SBC.
Then $Y = \phi(X)$ is a shift space over $A$.

Reason why

- It is not restrictive that $\phi$ is a 1-block code induced by $\Phi$.
- Put $\mathcal{L} = \{\Phi(w) | w \in \mathcal{B}(X)\}$. Clearly $\phi(X) \subseteq X_{A^* \setminus \mathcal{L}}$.
- Let $y \in X_{A^* \setminus \mathcal{L}}$. Then $y_{[-n,n]} = \Phi(x_{[-n,n]}^{(n)})$ for some $x_{[-n,n]}^{(n)} \in X$.
- Since $\mathcal{B}_{2k+1}(X)$ is finite for every $k$, a single $x \in X$ can be constructed s.t. $y_{[-n,n]} = \Phi(x_{[-n,n]})$ for every $n$.
- Then $y = \phi(x)$. 
Interlude: How to extract $x$ from the $x^{(n)}$’s

1. Take an infinite $S_0 \subseteq \mathbb{N}$ s.t. $x^{(n)}_0 = x^{(n')}_0$ for every $n, n' \in S_0$.
2. Take an infinite $S_1 \subseteq S_0$ s.t. $x^{(n)}_{[-1,1]} = x^{(n')}_{[-1,1]}$ for every $n, n' \in S_1$.
3. Take an infinite $S_2 \subseteq S_1$ s.t. $x^{(n)}_{[-2,2]} = x^{(n')}_{[-2,2]}$ for every $n, n' \in S_2$.
4. . . and so on, and so on . . .
5. Then

$$x_i = x^{(n)}_i \text{ for } n \in S_{|ij|}$$

is well defined.
The converse of a bijective SBC is a SBC

**Theorem**

Let $X$ be a subshift over $A$, $Y$ a subshift over $\mathcal{A}$.

Let $\phi : X \rightarrow Y$ be a bijective SBC.

Then $\phi^{-1} = \Psi_{\infty}^{[-N,N]}$ for some $N \geq 0$ and $\Psi : B_{2N+1}(Y) \rightarrow A$.

**Reason why**

- Again, it is not restrictive that $\phi$ is a 1-SBC.
- Suppose $\phi^{-1}(y)_0$ is not a function of $y_{[-n,n]}$ whatever $n$ is.
- Then, for every $n$, there are $x^{(n)}, \tilde{x}^{(n)} \in X$ s.t. $x_0^{(n)} \neq \tilde{x}_0^{(n)}$ but $\Phi(x^{(n)})_{[-n,n]} = \Phi(\tilde{x}^{(n)})_{[-n,n]}$.
- Similar to the previous theorem, $x \neq \tilde{x}$ can be found s.t. $\Phi(x)_{[-n,n]} = \Phi(\tilde{x})_{[-n,n]}$ for every $n \in \mathbb{N}$.
- Then $\phi(x) = \phi(\tilde{x})$, against bijectivity of $\phi$. 

Silvio Capobianco (Institute of Cybernetics)
April 14, 2010 27 / 34
A parallel with coding theory

In symbolic dynamics

- A \textit{subshift} is a special \textit{subspace} of a full shift.
- A \textit{code} is a special \textit{map} between subshifts.

In coding theory

- A \textit{code} is a special \textit{submonoid} of a free monoid.
- An \textit{encoder} is a special \textit{map} between codes.
Laurent series and polynomials

- A Laurent series on a field $\mathbb{F}$ is an expression

$$f(t) = \sum_{i=-\infty}^{+\infty} a_i t^i = \sum_{i=-\infty}^{+\infty} (f)_i t^i$$

with $a_i \in \mathbb{F}$ for all $i \in \mathbb{Z}$.

- A Laurent polynomial is a Laurent series where only finitely many $a_i$'s are non-zero.

- Laurent series can be multiplied by Laurent polynomials through

$$ (f \cdot g)_i = \sum_{j=-\infty}^{+\infty} (f)_j (g)_{i-j} $$
Convolutional encoders and codes

- Let $\mathbb{F}$ be a finite field.
- Identify the Laurent series $\sum_i a_i t^i$ with coefficients in $\mathbb{F}$ with the bi-infinite word $\ldots a_{-1} a_0 a_1 \ldots$ over $\mathbb{F}$.
- Let $G(t) = [g_{i,j}(t)]$ be a $k \times n$ matrix where each $g_{i,j}(t)$ is a Laurent polynomial over $\mathbb{F}$.
- A $(k, n)$-convolutional encoder is a transformation from the full $\mathbb{F}^k$-shift to the full $\mathbb{F}^n$-shift of the form

$$O(t) = E(I(t)) = I(t) \cdot G(t)$$

where the elements of $I(t)$ and $O(t)$ are Laurent series over $\mathbb{F}$.
- A $(k, n)$-convolutional code is the image of a convolutional encoder.
Example

Let \( I(t) = [I_1(t), I_2(t)] \) and

\[
G(t) = \begin{bmatrix} 1 & 0 & 1 + t \\ 0 & t & t \end{bmatrix}
\]

Then

\[
O(t) = [I_1(t), tl_2(t), (1 + t)I_1(t) + tl_2(t)]
\]

so that

\[
(O)_i = [(I_1)_i, (I_2)_{i-1}, (I_1)_i + (I_1)_{i-1} + (I_2)_{i-1}]
\]
From convolutions to sliding blocks

1. Let $O(t) = E(I(t)) = I(t) \cdot G(t)$ be a $(k, n)$-convolutional encoder.
2. Let $M$ and $N$ be the maximum and minimum power of $t$ in $G(t)$.
3. Identify the array of Laurent series over $\mathbb{F}$

$$[S_1(t), \ldots, S_r(t)]$$

with the bi-infinite word over $\mathbb{F}^r$

$$\ldots [(S_1)_{-1}, \ldots, (S_r)_{-1}][(S_1)_0, \ldots, (S_r)_0][(S_1)_1, \ldots, (S_r)_1] \ldots$$

4. Then $E = \Phi_{\infty}^{[-M,N]}$ with

$$(\Phi((I)_{-M} \ldots (I)_N))_s = \sum_{j=-M}^{N} \sum_{i=1}^{k} (I_i)_j ((G)_{i,s})_{-j}$$
And there is more...

Convolutional encoders are linear SBC
- Dependence of $O(t)$ from $I(t)$ is given by a set of linear equations.

Convolutional codes are linear irreducible subshifts
- Images of a full shift under a SBC.
- Subspaces of the (infinite-dimensional) $\mathbb{F}$-vector space $(\mathbb{F}^n)^\mathbb{Z}$ through a linear application.
- It is always possible to join $u$ and $v$ through a long enough $w$.

The converse also holds
There is a one-to-one correspondence between:
- Linear SBC and convolutional encoders.
- Linear irreducible subshifts and convolutional codes.
... and there shall be more...

- Shifts of finite type.
- Graphs and their shifts.
- Graphs as representations of shifts of finite type.
- State splitting.
- Shifts of finite type and data storage.

Thank you for attention!