I will review a relatively new approach to structural complexity that defines a process’s causal states and gives a procedure for finding them. It turns out that the causal-state representation—an $\epsilon$-machine—is the minimal one consistent with accurate prediction. The claim that this representation captures all of a process’s structure derives from $\epsilon$-machine optimality and uniqueness and on how $\epsilon$-machines compare to alternative representations. For example, one can directly relate measures of randomness and structural complexity obtained from $\epsilon$-machines to more familiar ones from ergodic and information theories.

These notes show the basics of $\epsilon$-machine reconstruction, illustrating the “batch” algorithm on a number of examples, both finite and infinite state.
Agenda

- Basics
- $\epsilon$-Machine Reconstruction
- Entropy and Complexity
- Examples for Everyone
- The Wild Side
Processes

Observations $\rightarrow S = \leftarrow S \rightarrow S$

$past \rightarrow future$

$\ldots s_{-L}s_{-L+1}\ldots s_{-1}s_{0}|s_{1}\ldots s_{L-1}s_{L}\ldots$, $s_{i} \in A$

Entropy rate (degree of unpredictability)

$h_{\mu} = H(Pr(s_{i}|s_{i-1}s_{i-2}s_{i-3}\ldots))$

How much past information transmitted to future? (Excess Entropy)

$E = H\left(\frac{Pr(\rightarrow S)}{Pr(\leftarrow S)Pr(\rightarrow S)}\right)$

*Note*: Property that describes “raw” sequences.
Prediction Game
Causal States and $\epsilon$-Machines

**Morph** = Set of possible (future) behaviors, given knowledge of past

**Causal State** = Same future morph at different times

Causal state = identical conditions of ignorance/knowledge

---

**Probabilistic Morph** = $\Pr(\text{Future}|\text{Past})$

**Causal State** = Equivalence class of probabilistic morphs

$t_i \sim t_j \Leftrightarrow \Pr(\text{Future}|\text{Past at } t_i) = \Pr(\text{Future}|\text{Past at } t_j)$

Set of causal states: $S = \{\text{Classes of histories induced by relation}\$

Transition dynamic: $\mathcal{T} : S \rightarrow S$

$\epsilon$-Machine = $\{S, \mathcal{T}\$

Always unique start state: $S_0 \in S$

Entropy rate (degree of unpredictability)

$$h_\mu = H\left[\Pr\left(S \rightarrow S'|S\right)\right]$$

Statistical Complexity (size of $\epsilon$-machine)

$$C_\mu = H[\Pr(S)]$$
Mathematical Foundations

 Lemma 1: Given $\mathcal{S}$, past and future are conditionally independent.

 Proposition: $\epsilon$-Machines are Semi-Groups.

 Lemma 2: $\epsilon$-Machines are Deterministic.

 Definition: $\epsilon$-Machine Reconstruction

\[ \Pr(\vec{s}) \rightarrow \{\mathcal{S}, \mathcal{T}\}. \]

 Theorem 1: Causal States are Maximally Predictive,

\[ H[\vec{s}^L|\mathcal{R}] \geq H[\vec{s}^L|\mathcal{S}] , \text{ where } \mathcal{R} = \text{ rival states.} \]

 Corollary: Causal States are Sufficient Statistics.

 Theorem 2: Causal States are Minimal,

\[ C_\mu(\hat{\mathcal{R}}) \geq C_\mu(\mathcal{S}) , \text{ where } \hat{\mathcal{R}} = \text{ prescient rivals.} \]

 Theorem 3: Causal States are Unique,

\[ C_\mu(\hat{\mathcal{R}}) = C_\mu(\mathcal{S}) \Rightarrow \hat{\mathcal{R}} = f(\mathcal{S}) \land \mathcal{S} = g(\hat{\mathcal{R}}). \]

 Theorem 4: $\epsilon$-Machines are Minimally Stochastic,

\[ H[\hat{\mathcal{R}}'|\hat{\mathcal{R}}] \geq H[\mathcal{S}'|\mathcal{S}]. \]

 Theorem 5: Statistical Complexity upper bounds Excess Entropy,

\[ \mathbf{E} \leq C_\mu(\mathcal{S}). \]

1. $\mathbf{E} \neq \text{“stored information”}; C_\mu$ is.
2. $\mathbf{E}$ is apparent information ... from sequences, not internal states.
3. Theorem 5 $\Rightarrow$ we must build models.
$\epsilon$-Machine Reconstruction

History: Geometry from a Time Series

• Steps
  a. Parse Tree of depth D
  b. Morphs of depth L
  c. Causal States
  d. Causal Transitions
  e. Symbol Transition Matrices
  f. Connection Matrix

• Topological reconstruction, first

• Examples
  a. Constant Process: $1^*$
  b. Fair Coin: $(0 + 1)^*$
  c. Period-2: $(10)^*$

• Comments
  a. Complexity-Entropy Diagram
  c. Recurrent v. transient states

• Probabilistic reconstruction
Randomness and Structure: Entropy v. Complexity

- Topological entropy:
  \[ h_0 = \log_2 \lambda \]

- Topological complexity:
  \[ C_0 = \log_2 |\mathcal{S}| \]

- Entropy rate:
  \[ h_\mu = - \sum_{v \in \mathcal{S}} p_v \sum_{s \in \mathcal{A}} p_{v \rightarrow v'} s \log_2 p_{v \rightarrow v'} s \]

- Statistical complexity:
  \[ C_\mu = - \sum_{v \in \mathcal{S}} p_v \log_2 p_v \]

- Eigenvalues and eigenvectors.
- For above examples.
Examples for All

• A Period-3 Process: (101)*.
• Golden Mean Process: No consecutive 0s.
• Noisy Period-2 Process: (1Σ)*, where Σ = {0, 1}.
• Even Process: 1s occur in blocks of even number, bordered by 0s.

What to See and Do

• Reconstruct ε-Machine
  a. Parse Tree and Morphs
  b. State-Transition Diagram
  c. Symbol Transition Matrices
• Calculate entropy rate and statistical complexity
• Interpret ε-machine structure: What do the states mean in each case?
The Wild Side

• Morse-Thue Sequence (Onset of Chaos via Period-Doubling): $0 \rightarrow 11$ and $1 \rightarrow 10$.
• Simple Nondeterministic Source.
• More General Hidden Markov Processes.