Memory in Processes II

Reading for this lecture:

CMR article RURO.
Memory in Processes II ...

Classes of Excess Entropy:

**Finitary process**: \( E < \infty \)

Exponential or finite-length convergence

**Infinitary process**: \( E \to \infty \)

Examples:

\( \infty \)-order Markov: Even Process

Topological complexity: Morse-Thue Process

Stochastic complexity: Simple Nonunifilar Source
Memory in Processes II ...  
Classes of Excess Entropy ... 

Even Process: After pair of 1s, coin flip

Presentation as Unifilar HMM

Even Process: After pair of 1s, coin flip

No finite-order Markov process exactly models the Even process.

But,

$$E \approx 0.902 \text{ bits}$$
Memory in Processes II ...

Classes of Excess Entropy ...

Even Process ...

∞-order Markov process.

But, still exponential entropy-rate decay:

\[ h_\mu(L) - h_\mu \propto 2^{-\gamma L} \]

\[ \gamma \approx \frac{1}{2} \]

Lecture 17: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Morse-Thue Process:
Support is a context-free language
Generated by Logistic map at onset of chaos

Production rules:
\[ \sigma(0) = 01 \]
\[ \sigma(1) = 10 \]

For example:
\[ \sigma^5(1) = 10010110011010010110100110010110 \]

Aperiodic, infinite memory, predictable!
Classes of Excess Entropy ...

Exact entropy-rate approximates:

\[ h_\mu(1) = 1 \]

\[ h_\mu(2) = \log_2 3 - \frac{2}{3} \]

\[ h_\mu(3) = \frac{2}{3} \]

\[ h_\mu(L) = \begin{cases} 
\frac{4}{3 \cdot 2^k}, & \text{if } 2^k + 1 \leq L - 1 \leq 3 \cdot 2^{k-1} \\
\frac{2}{3 \cdot 2^k}, & \text{if } 3 \cdot 2^{k-1} + 1 \leq L - 1 \leq 2^{k+1} 
\end{cases} \]

Slow entropy convergence (power-law):

\[ h_\mu(L) \propto \frac{1}{L} \]

Entropy-rate vanishes:

\[ h_\mu = 0 \text{ bits per symbol} \]

\[ H(L) \propto \log_2(L) \]
Memory in Processes II ...
Classes of Excess Entropy ...

Excess entropy diverges:
Arbitrarily long-range correlations
(e.g., critical phenomena at phase transitions)

Infinitary Process!
$E \to \infty$

$E(L) = H(L) - h_\mu L$
$E'(L) = I(\overrightarrow{X^{L/2}}; \overrightarrow{X^{L/2}})$

Lecture 17: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Simple Nonunifilar Source:

What is its entropy rate?
Recall: Cannot use nonunifilar presentation to answer.
Memory in Processes II ...
Classes of Excess Entropy ...
Simple Nonunifilar Source ...

Entropy curves
Memory in Processes II ...

Classes of Excess Entropy ...

Simple Nonunifilar Source ...

∞-order Markov process.

Neither exponential decay:

\[ h_\mu(L) - h_\mu \propto 2^{-\gamma L} \]

Nor power-law decay:

\[ h_\mu(L) - h_\mu \propto L^\alpha \]
Synchronization:

Problem Statement:
You have a correct model of a process, but you don’t know it’s current state.

Question: How much information must you extract from measurements to know which hidden state the process is in?
Transient Information:

Synchronized to source when:

\[ L \geq L' \]

you have

\[ H(L) \approx E + h_\mu L \]

Synchronized:

At length \( L' \) at which you see true entropy rate.

Extracted sufficient information to do optimal prediction.

Lecture 18: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Memory in Processes ...
Transient Information ...

How much information to extract?

Transient Information:

\[ T = \sum_{L=0}^{\infty} \left[ E + h \mu L - H(L) \right] \]

Controls convergence to synchronization. Units: bits x symbols
Example of Transient Information:

Tahitian Vacation (3 days)!

Weather has a 5 day cycle:
  Two days of rain, followed by three of sun

Weather is exactly predictable: $h_\mu = 0$ bits per day

Weather has memory: $E = \log_2 5$ bits

But,

How to pack?
What to pack?
What to wear on trip?
Dressed appropriately for arrival?
Memory in Processes ...
Example of Transient Information ...
Tahitian Vacation ... packing

No weather reports yet.

0 = Rain
1 = Sun

Weather Reports
Update Traveler's Model

Tahiti

No weather reports yet.
Memory in Processes ...
Example of Transient Information ...
Tahitian Vacation ... packing

No weather reports yet.

Rain!

Tahiti
Weather Reports
Update Traveler's Model

Lecture 18: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Memory in Processes...

Example of Transient Information...

Tahitian Vacation...packing

No weather reports yet.

Rain!

Sun!

Lecture 18: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Memory in Processes ...
Example of Transient Information ...
Tahitian Vacation ... packing

No weather reports yet.

Rain!

Sun!

Pack umbrella, wear shorts on plane

\[ T \approx 4.073 \text{ bit } \times \text{symbols} \]

Lecture 18: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Memory in Processes ...  
Transient Information ...

How to interpret?

\[ H(L) \]

\[ \mathbb{E} + h_{\mu}L \]

\[ H(L) \]

Lecture 18: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
Synchronization information:

Observer has correct model of a Markov chain: $\mathcal{M} = \{V, T\}$

Observer Synchronized to Process:

$$T(L) \equiv E + h_\mu L - H(L) = 0$$

Observer knows with certainty in which state the process is:

$$\Pr(v_0, v_1, \ldots, v_k) = (0, \ldots, 1, \ldots, 0)$$

Average per-symbol uncertainty is exactly $h_\mu$. 
Synchronization information:

\[
\mathcal{S} \equiv \sum_{L=0}^{\infty} \mathcal{H}(L)
\]
Memory in Processes ...

Transient Information ...

Synchronization information ...

Theorem: For a R-block (spin-block) process, the synchronization information is given by:

\[ S = T + \frac{1}{2} R(R + 1) h_\mu \]

Corollary: For periodic process:

\[ S = T \]
Memory in Processes ...
Transient Information ...

How to interpret?

1. Total uncertainty observed while synchronizing.

2. Information to extract to be synchronized.
Examples of Transient Information:

Fair & Biased Coins

& IID Processes: $T = 0$

![Graph showing $H(L)$ for fair and biased coins](image)

$L$: Length of sequence

$H(L)$: Entropy of the sequence

- $H(L)$: Fair Coin
- $H(L)$: Biased Coin, $p=0.7$
Memory in Processes ...
Examples of Transient Information ...

Period-5 Processes:

There are three distinct:

\[(11000)\infty\]
\[(10101)\infty\]
\[(10000)\infty\]

All:

Predictable: \(h_\mu = 0 \text{ bits}\)
Memory: \(E = \log_2 5 \text{ bits}\)
Memory in Processes ...
Examples of Transient Information ...

Period-5 Processes ...

But different ways to sync:
Memory in Processes ...
Examples of Transient Information ...

Period-5 Processes ...

But different ways to sync:

\((11000)^\infty\) \(T \approx 4.073\) bit \(\times\) symbols
\((10101)^\infty\) \(T \approx 4.873\) bit \(\times\) symbols
\((10000)^\infty\) \(T \approx 5.273\) bit \(\times\) symbols

\[\begin{array}{|c|}
11000: T = 4.073 \\
10101: T = 4.873 \\
10000: T = 5.273 \\
\end{array}\]
Memory in Processes ...
Examples of Transient Information ...

Period-P Processes:

Entropy rate vanishes.

Excess entropy same for all.

But T distinguishes periodic processes.
Memory in Processes ...

Transient Information For Periodic Processes

Period: $P$

Max $T$:
Slow convergence
Most nonuniform word dist.
P−1 0’s followed by isolated 1

$T_{\text{max}} \approx \frac{1}{2} P \log_2 P$

Min $T$:
Fast convergence
Flattest word distribution
$P = 2^L$: deBruin sequences

$T_{\text{min}} \approx \frac{1}{2} \left( \log_2^2 P + \log_2 P \right)$
Information-Entropy Roadmap for a Stochastic Process:

\[ H(L) \]

\[ E + h_\mu L \]

\[ H(L) \]

\[ h_\mu L \]

\[ 0 \]

\[ L \]
Memory in Processes ...

Regularities Unseen, Randomness Observed:
• Untangle distinct sources of apparent randomness?
• Estimates of entropy rate if ignore a process’s structure?

Consequences:
• When an observer ignores entropy-rate convergence?
• When the process’s apparent memory is ignored?
• If the observer ignores synchronization?
• If the observer assumes it is synchronized?
Memory in Processes ...

Disorder is the Price of Ignorance:

Ignore process’s memory

By assuming

\[ E = 0 \]

Over-estimate true randomness

\[ h_\mu' > h_\mu \]

Lesson:

Structure (E & T) converted to apparent randomness \( (h_\mu) \).
Predictability and Instantaneous Synchronization:

Instant Sync:
Assume you know memory $E$

Your estimate $\hat{h}_\mu$ of unpredictability?

$\hat{h}_\mu < h_\mu$

Lesson:
Assumed synchronization converted to false predictability.
Memory in Processes ...

Assumed Synchronization Implies Reduced Apparent Memory:

Assume you’re sync’d:

\[ H(L) = E + h_L L \]
\[ h_L(L) = h_L \]

Estimate of memory?

\[ \hat{E} < E \]

Lesson:

The world appears less structured.

---

Lecture 18: Natural Computation & Self-Organization, Physics 256A (Winter 2014); Jim Crutchfield
**Calculus of the Entropy Hierarchy:**

Via Discrete-Time Derivatives and Integrals

<table>
<thead>
<tr>
<th>Level</th>
<th>Gain (Derivative)</th>
<th>Information (Integral)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Block Entropy $H(L)$</td>
<td>Transient Information $T = \sum_{L=1}^{\infty} [E + h_\mu L - H(L)]$</td>
</tr>
<tr>
<td>1</td>
<td>Entropy Rate Loss $h_\mu(L) = \Delta H(L)$</td>
<td>Excess Entropy $E = \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$</td>
</tr>
<tr>
<td>2</td>
<td>Predictability Gain $\Delta^2 H(L)$</td>
<td>Total Predictability (Redundancy) $G = -\mathcal{R}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Equations:

- $h_\mu(L) = \Delta H(L)$
- $\Delta^2 H(L)$
- $T = \sum_{L=1}^{\infty} [E + h_\mu L - H(L)]$
- $E = \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$
Memory in Processes ...

Reading for next lecture:

Yeung and Anatomy in CMech Reader.

Course Evaluations (15 minutes).