Memory in Processes

Memory in Processes ...

Entropy Hierarchy:

Take derivatives:

(1) Block entropy: $H(L)$

(2) Entropy rate: $h_\mu(L) = \Delta H(L)$

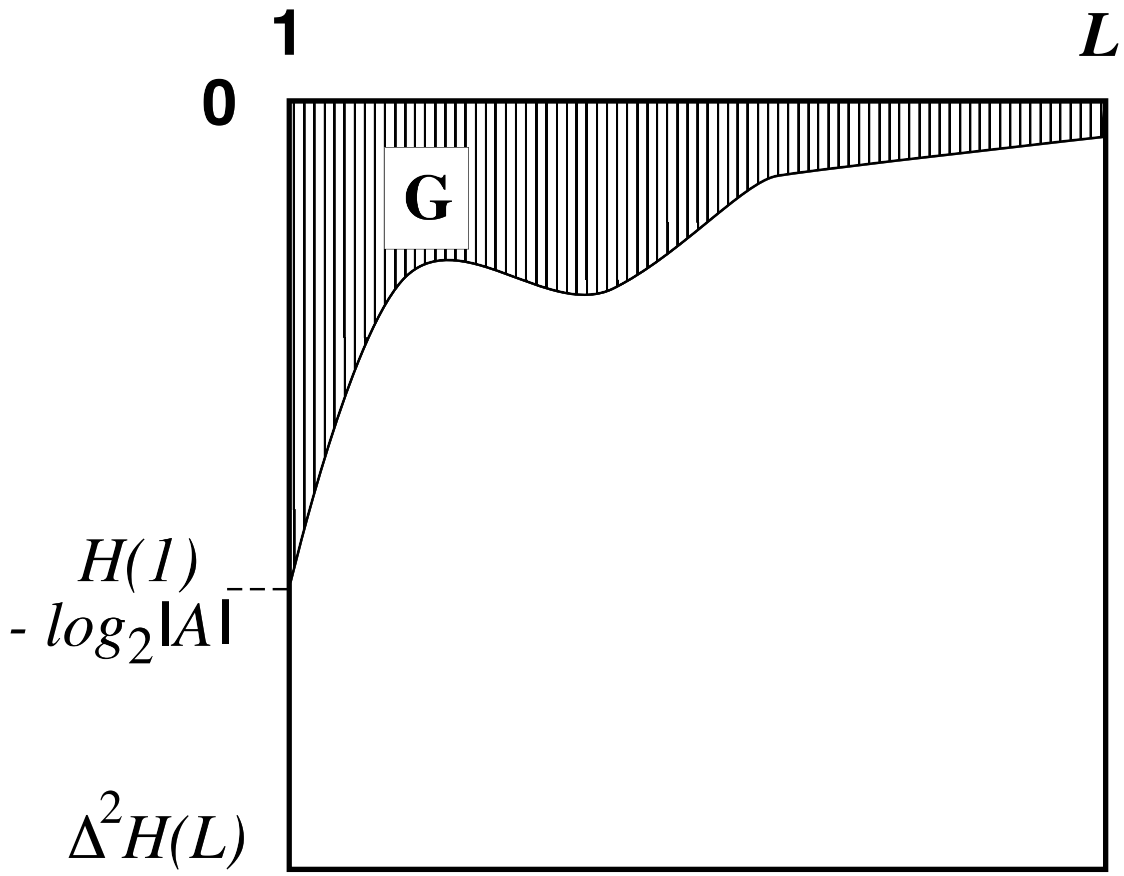
(3) Predictability gain: $\Delta h_\mu(L) = \Delta^2 H(L)$

Now take integrals!

Memory in Processes ...

Total Predictability:

$$\mathbf{G} = \sum_{L=1}^{\infty} \Delta^2 H(L)$$



Redundancy:

$$-\mathbf{G} = \mathcal{R} = \log_2 |\mathcal{A}| - h_\mu$$

Interpretation:

(1) Account for all correlations to see intrinsic randomness

(2) Until that point, correlations appear as *excess randomness*

Memory in Processes ...

Excess Entropy:

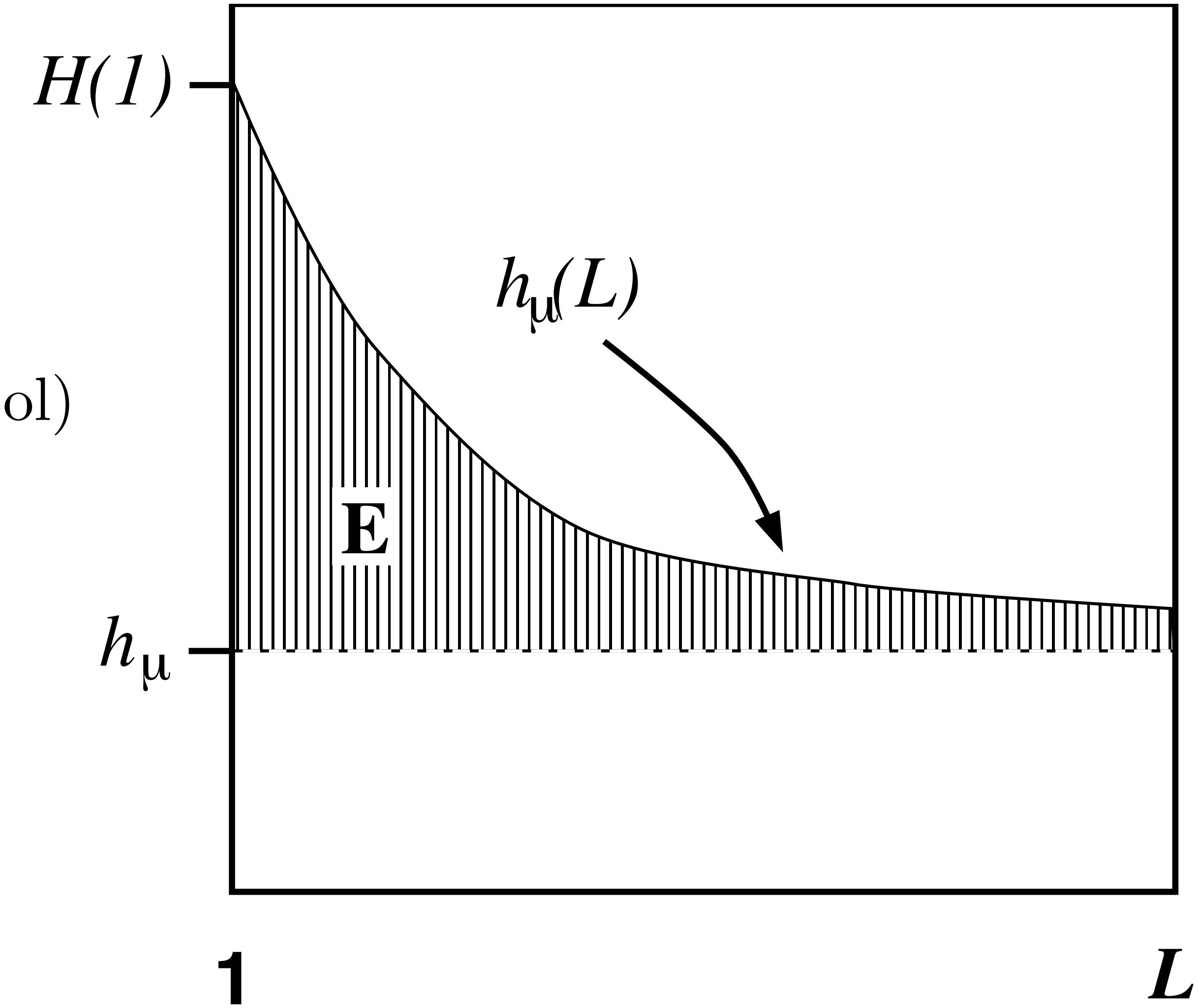
As entropy convergence:

$$\mathbf{E} = \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$$

$(\Delta L = 1 \text{ symbol})$

As intrinsic redundancy:

$$\mathbf{E} = \sum_{L=1}^{\infty} r(L)$$



Properties:

(1) Units: $\mathbf{E} = [\text{bits}]$

(2) Positive: $\mathbf{E} \geq 0$

(3) Controls convergence to actual randomness.

(4) Slow convergence ⇔ Correlations at longer words.

(5) Complementary to entropy rate.

Memory in Processes ...

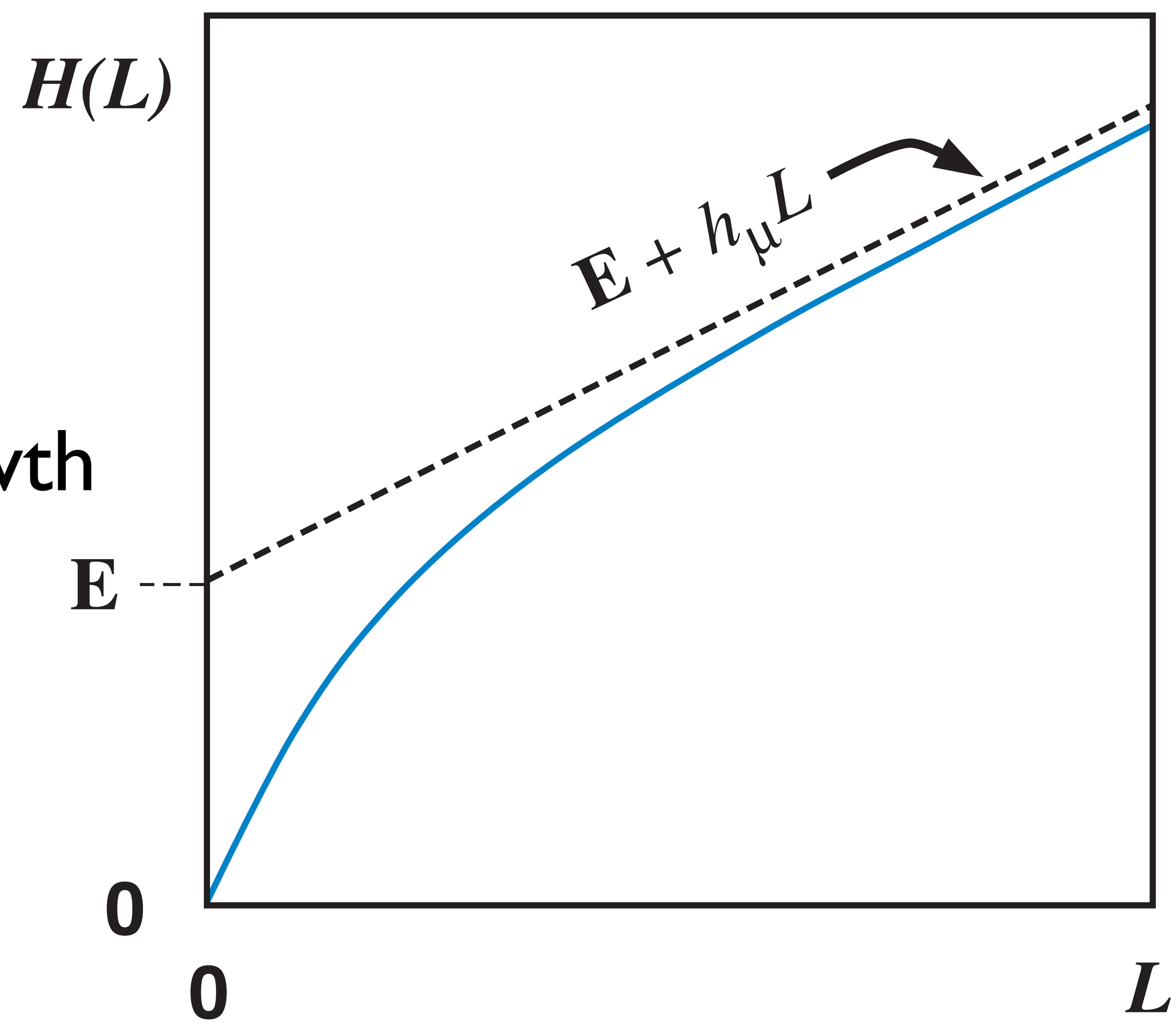
Excess Entropy ...

Asymptote of entropy growth:

$$\mathbf{E} = \lim_{L \to \infty} \left[H(L) - h_\mu L \right]$$

That is,

$$H(L) \propto \mathbf{E} + h_\mu L$$



Y-Intercept of entropy growth

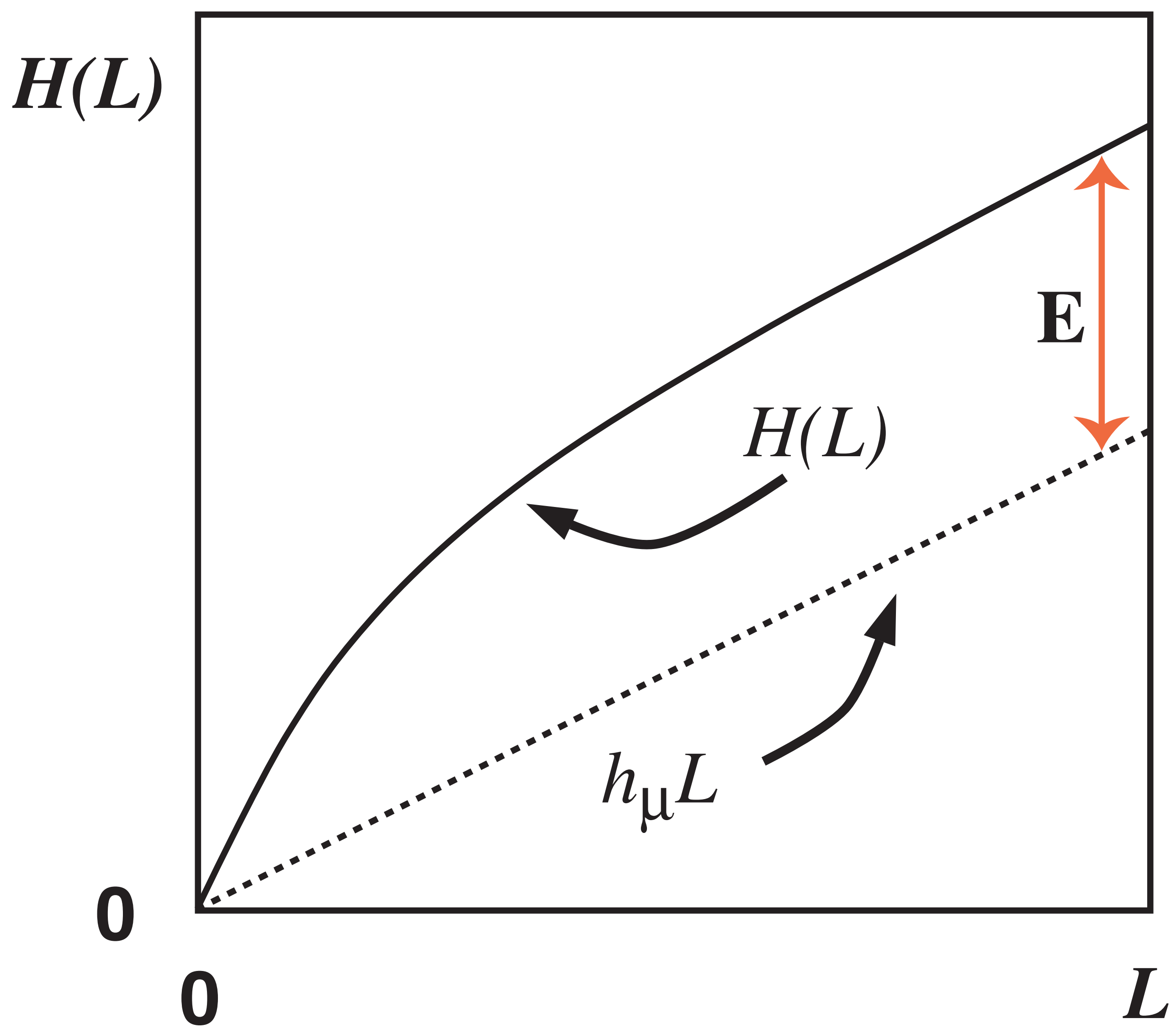
Memory in Processes ...

Excess Entropy ...

Cost of Amnesia:

Forget what you know:

Information needed to recover predicting with error ~h_μ



Cf. Memoryless Source: IID at same entropy rate

Memory in Processes ...

Excess Entropy ...

Mutual information between past and future:

View process as a communication channel: Past to Future

$$\mathbf{E} = I(\overleftarrow{S}; \overrightarrow{S})$$

Property:

Symmetric in time

Interpretation:

Information that process communicates from past to future.

Reduction in uncertainty about the future, given the past.

Reduction in uncertainty about the past, given the future.

Memory in Processes ...

Examples of Excess Entropy:

Fair Coin:

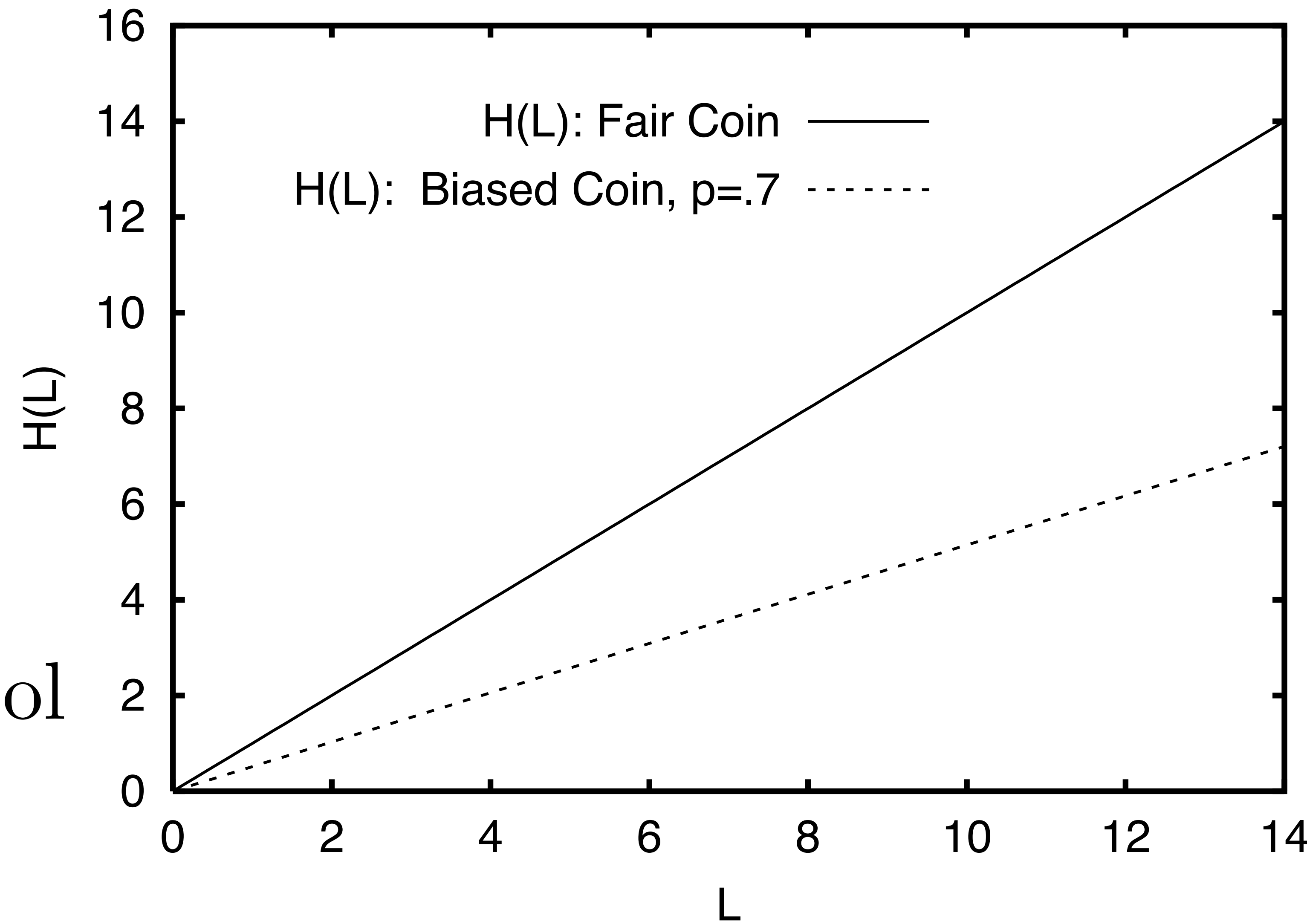
$$h_\mu = 1 \text{ bit per symbol}$$

$$\mathbf{E} = 0 \text{ bits}$$

Biased Coin:

$$h_\mu = H(p) \text{ bits per symbol}$$

$$\mathbf{E} = 0 \text{ bits}$$



Any IID Process:

$$h_\mu = H(X) \text{ bits per symbol}$$

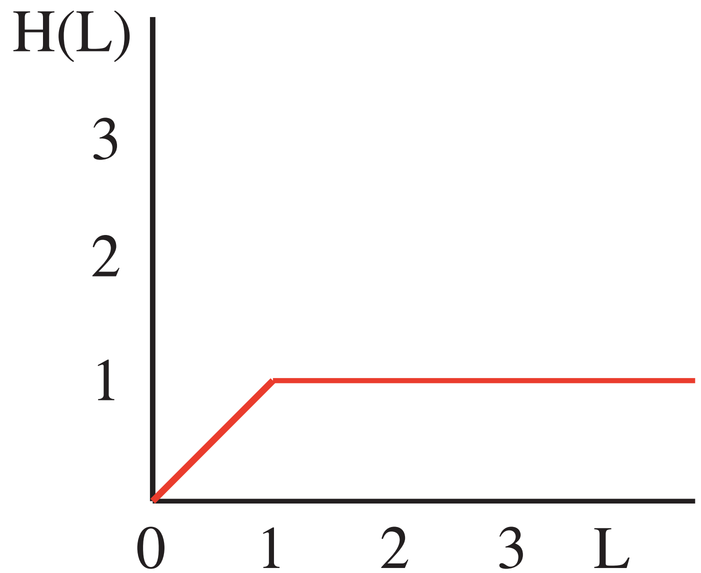
$$\mathbf{E} = 0 \text{ bits}$$

Memory in Processes ...

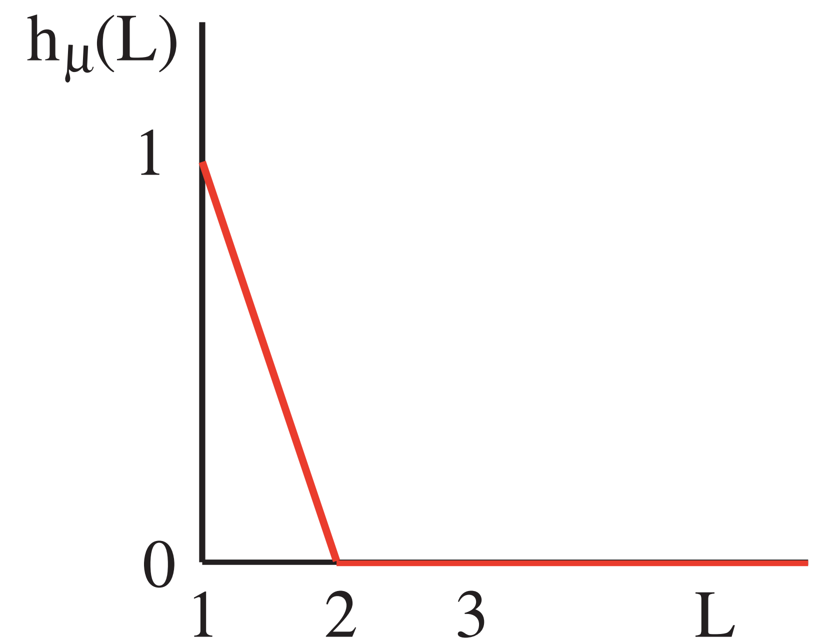
Examples of Excess Entropy ...

Period-2 Process: 010101010101

$$H(1) = 1$$
$$H(2) = 1$$
$$H(3) = 1$$



$$h_\mu(1) = 1$$
$$h_\mu(2) = 0$$
$$h_\mu(3) = 0$$



$$h_\mu = 0 \text{ bits per symbol}$$

$$\mathbf{E} = 1 \text{ bit}$$

Meaning:

1 bit of phase information

0-phase or 1-phase?

Memory in Processes ...

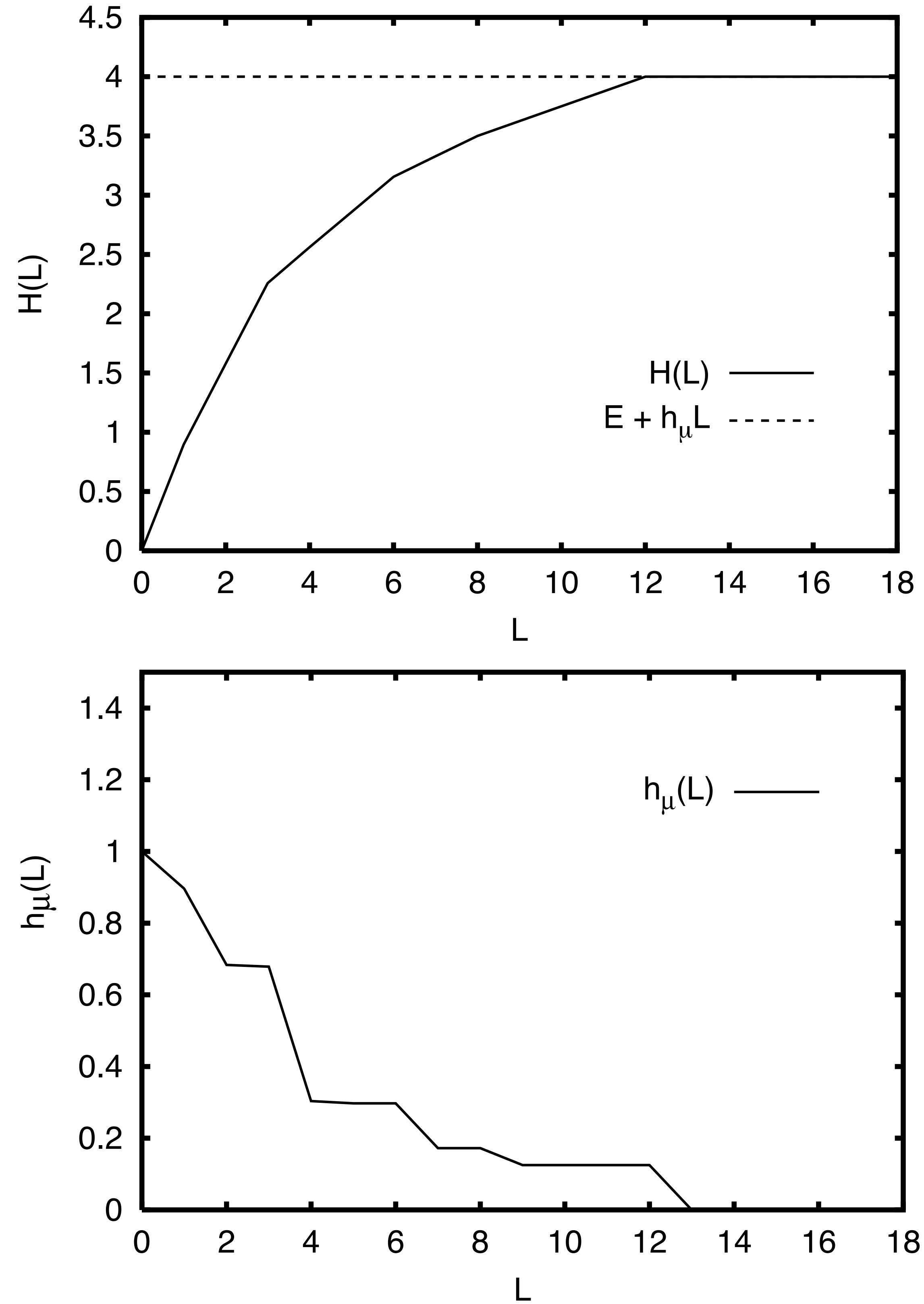
Examples of Excess Entropy ...

Period-16 Process:

$$(1010111011101110)^\infty$$

$$h_\mu = 0 \text{ bits per symbol}$$

$$\mathbf{E} = 4 \text{ bits}$$



Period-P Processes:

$$h_\mu = 0 \text{ bits per symbol}$$

$$\mathbf{E} = \log_2 P \text{ bits}$$

Cf., entropy rate does not distinguish periodic processes!

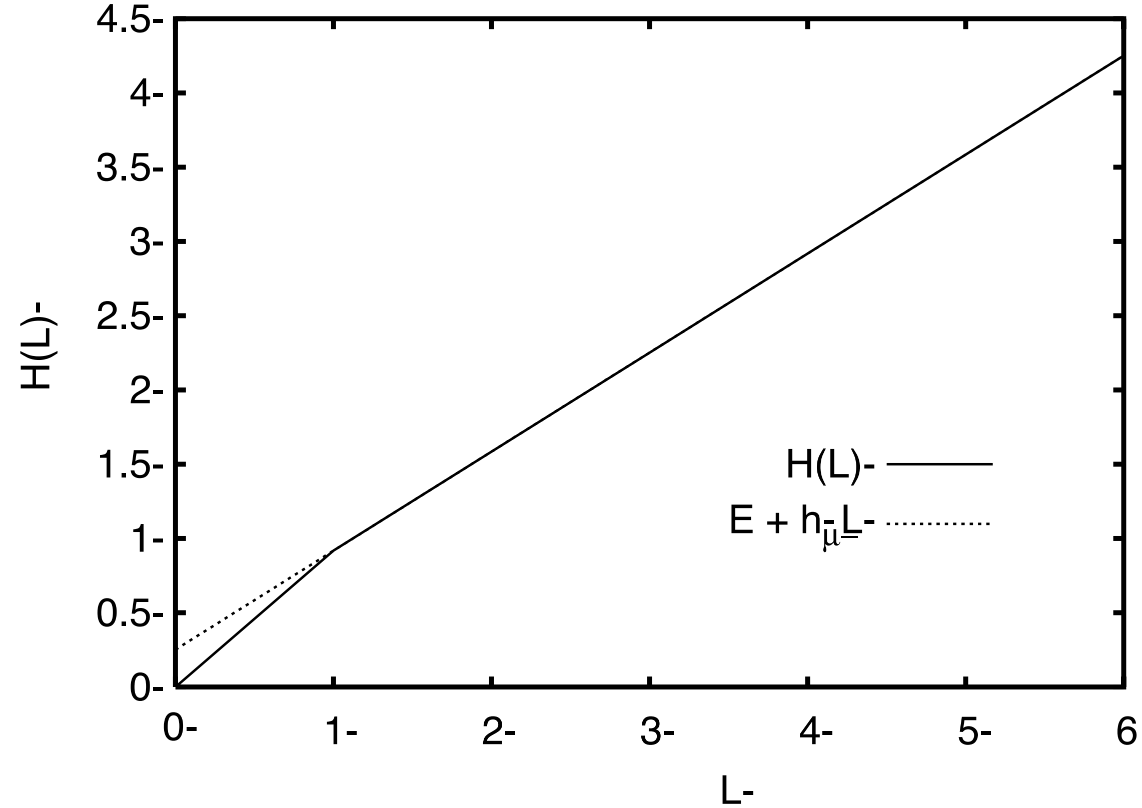
Memory in Processes ...

Examples of Excess Entropy ...

Golden Mean Process:

$$h_\mu = \frac{2}{3} \text{ bits per symbol}$$

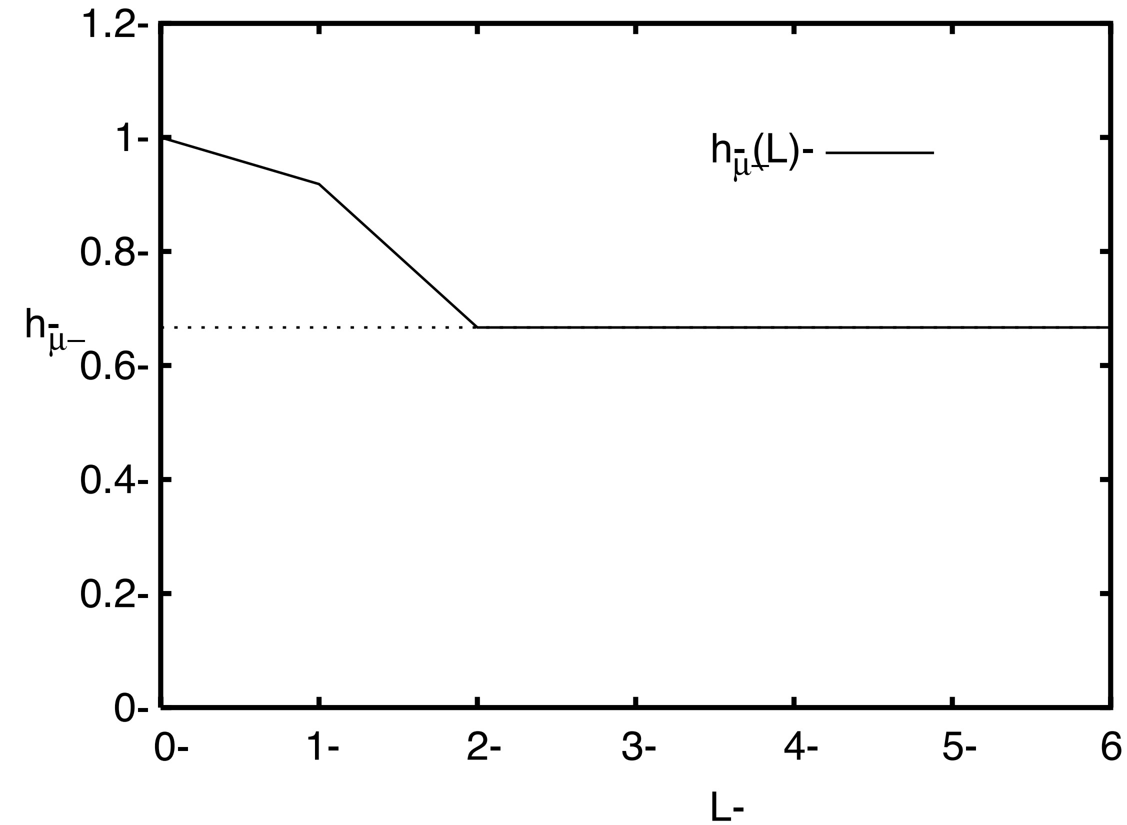
$$\mathbf{E} \approx 0.2516 \text{ bits}$$



R-Block Markov Chain:

$$\mathbf{E} = H(R) - R \cdot h_\mu$$

(E.g., 1D Ising Spin System)



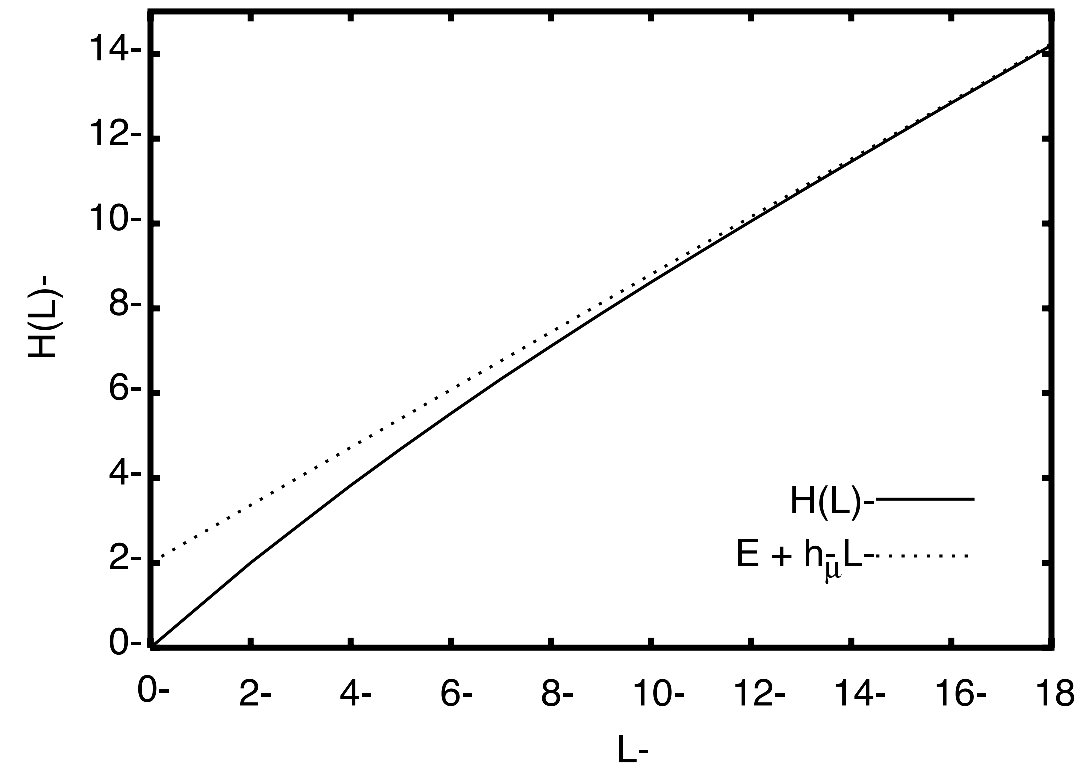
Memory in Processes ...
Examples of Excess Entropy:
Finitary Processes: Exponential entropy convergence

Random-Random
XOR (RRXOR) Process:

$$S_t = S_{t-1} \text{ XOR } S_{t-2}$$

$$h_\mu = \frac{2}{3} \text{ bits per symbol}$$

$$\mathbf{E} \approx 2.252 \text{ bits}$$

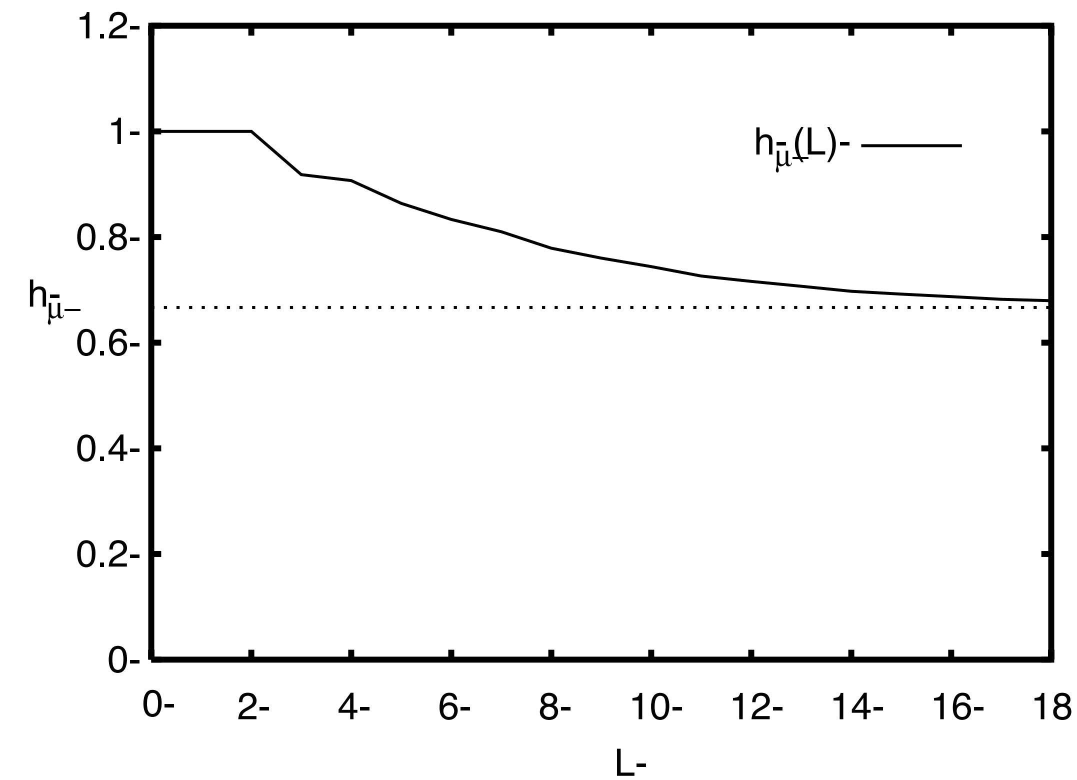


Finitary processes:
Exponential convergence:

$$h_\mu(L) - h_\mu \approx 2^{-\gamma L}$$

$$\mathbf{E} = \frac{H(1) - h_\mu}{1 - 2^{-\gamma}}$$

$$\gamma \approx 0.30$$



Memory in Processes ...

Examples of Excess Entropy:

Infinitary Processes:

$$\mathbf{E} \to \infty$$

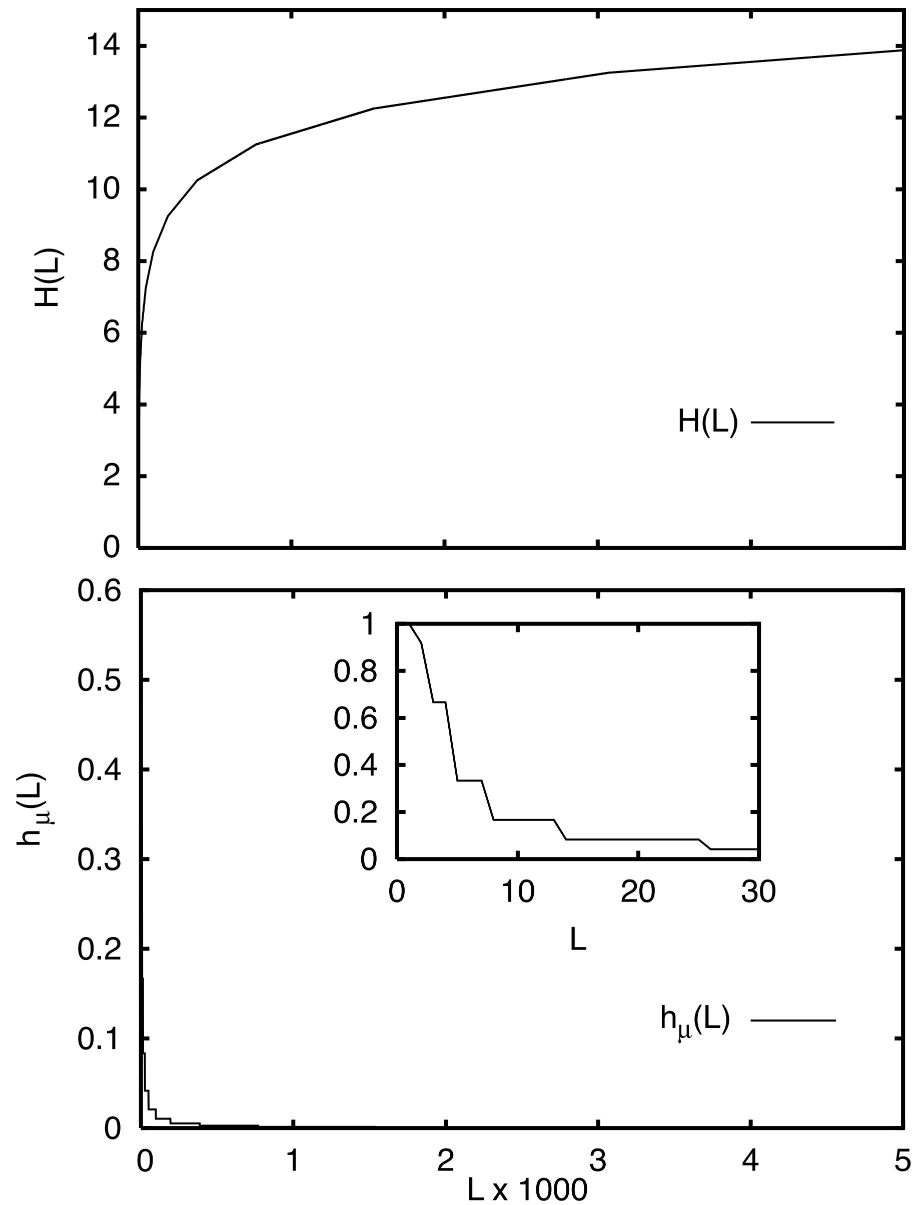
Excess entropy can diverge:

- Slow entropy convergence
- Long-range correlations
  - (e.g., at phase transitions)

Morse-Thue Process:

A context-free language

From Logistic map at onset of chaos



$$h_\mu = 0 \text{ bits per symbol}$$

Memory in Processes ...

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?

Memory in Processes ...

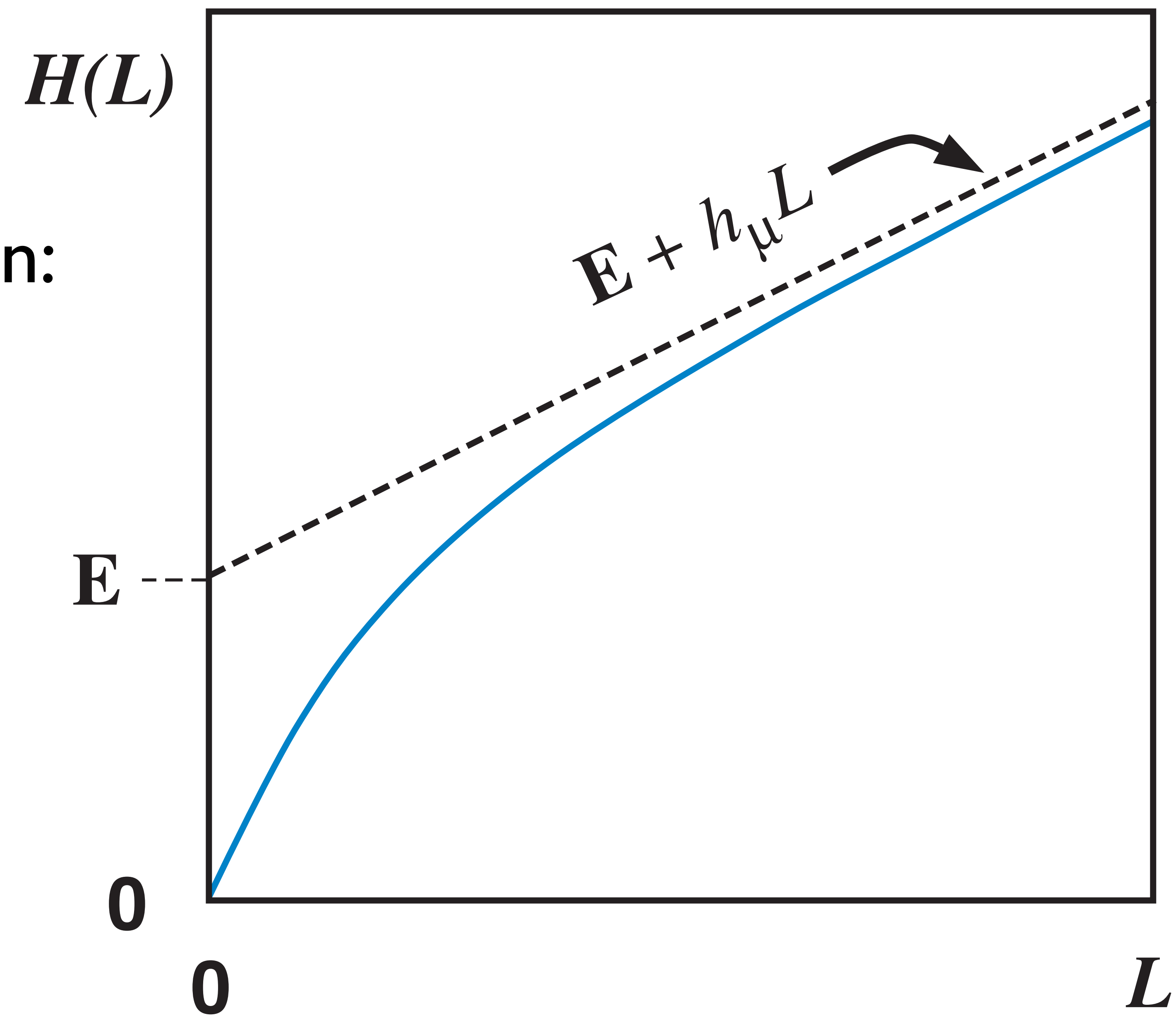
Transient Information:

Synchronized to source when:

$$L \geq L'$$

you have

$$H(L) \approx \mathbf{E} + h_\mu L$$



Synchronized:

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

Extracted sufficient information to do optimal prediction.

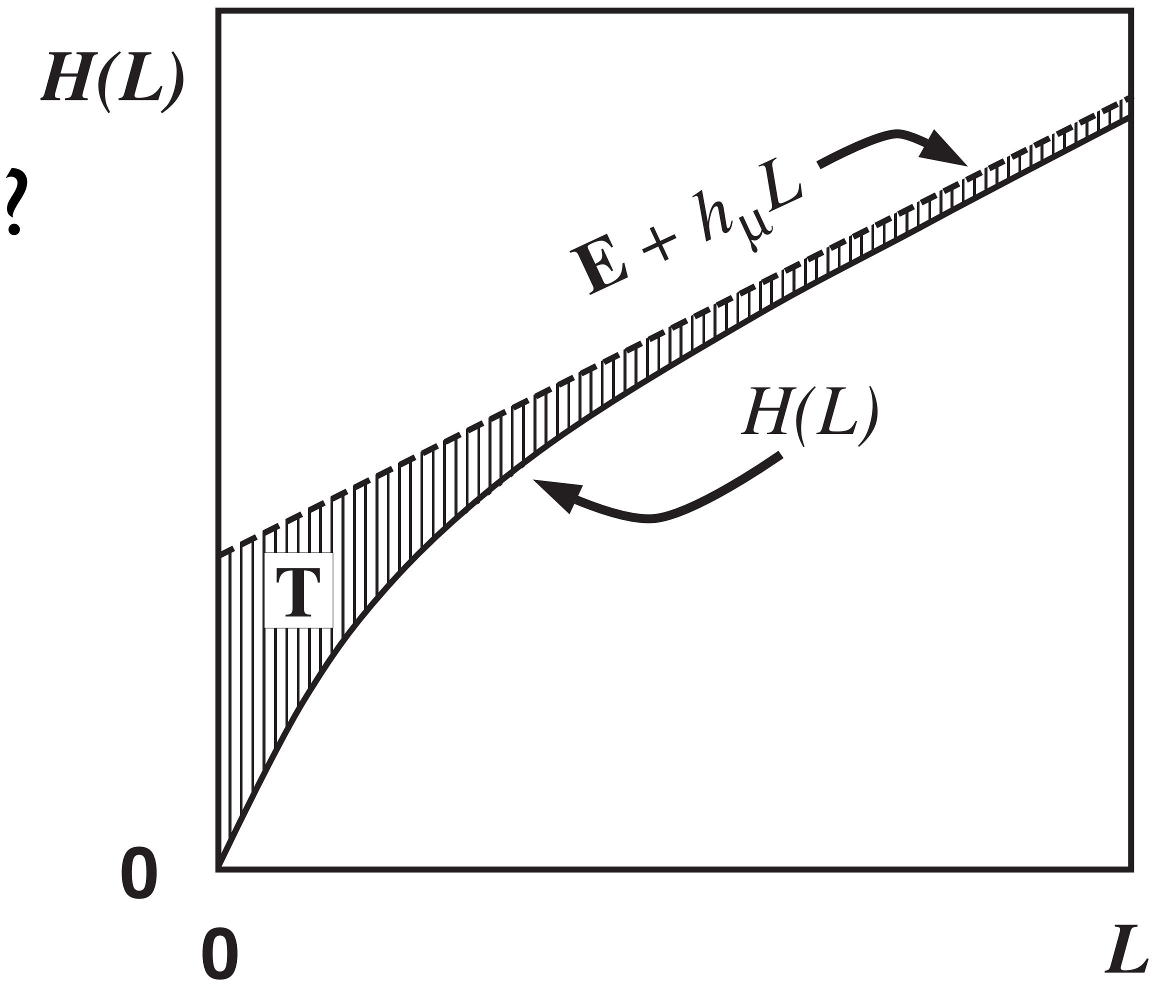
Memory in Processes ...

Transient Information ...

How much information to extract?

Transient Information:

$$\mathbf{T} = \sum_{L=0}^{\infty} [\mathbf{E} + h_\mu L - H(L)]$$



Controls convergence to synchronization.

Units: bits x symbols

Memory in Processes ...

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: $h_\mu = 0 \text{ bits per day}$

Weather has memory: $\mathbf{E} = \log_2 5 \text{ bits}$

But,

How to pack?

What to pack?

What to wear on trip?

Dressed appropriately for arrival?

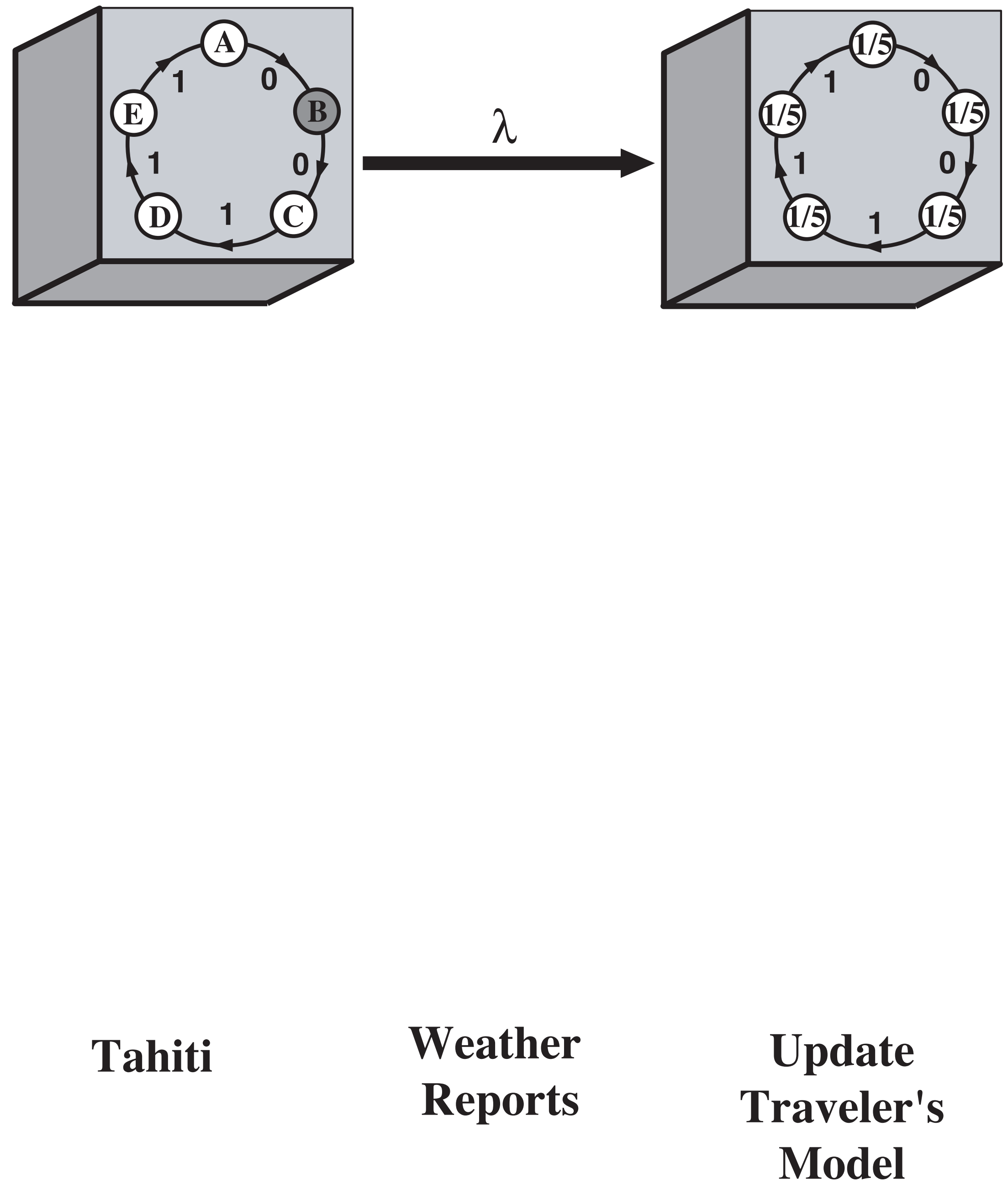
Memory in Processes ...

Example of Transient Information ...

Tahitian Vacation ... packing

0 = Rain
1 = Sun

No weather reports yet.



Memory in Processes ...

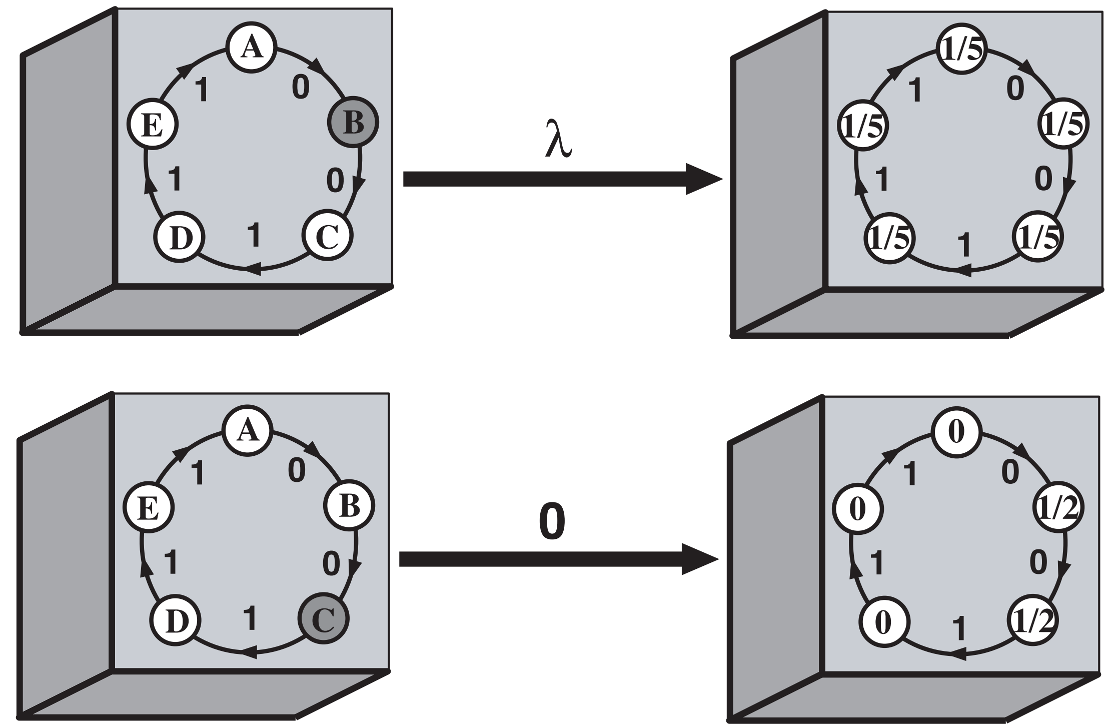
Example of Transient Information ...

Tahitian Vacation ... packing

0 = Rain
1 = Sun

No weather reports yet.

Rain!



Tahiti

Weather Reports

Update Traveler's Model

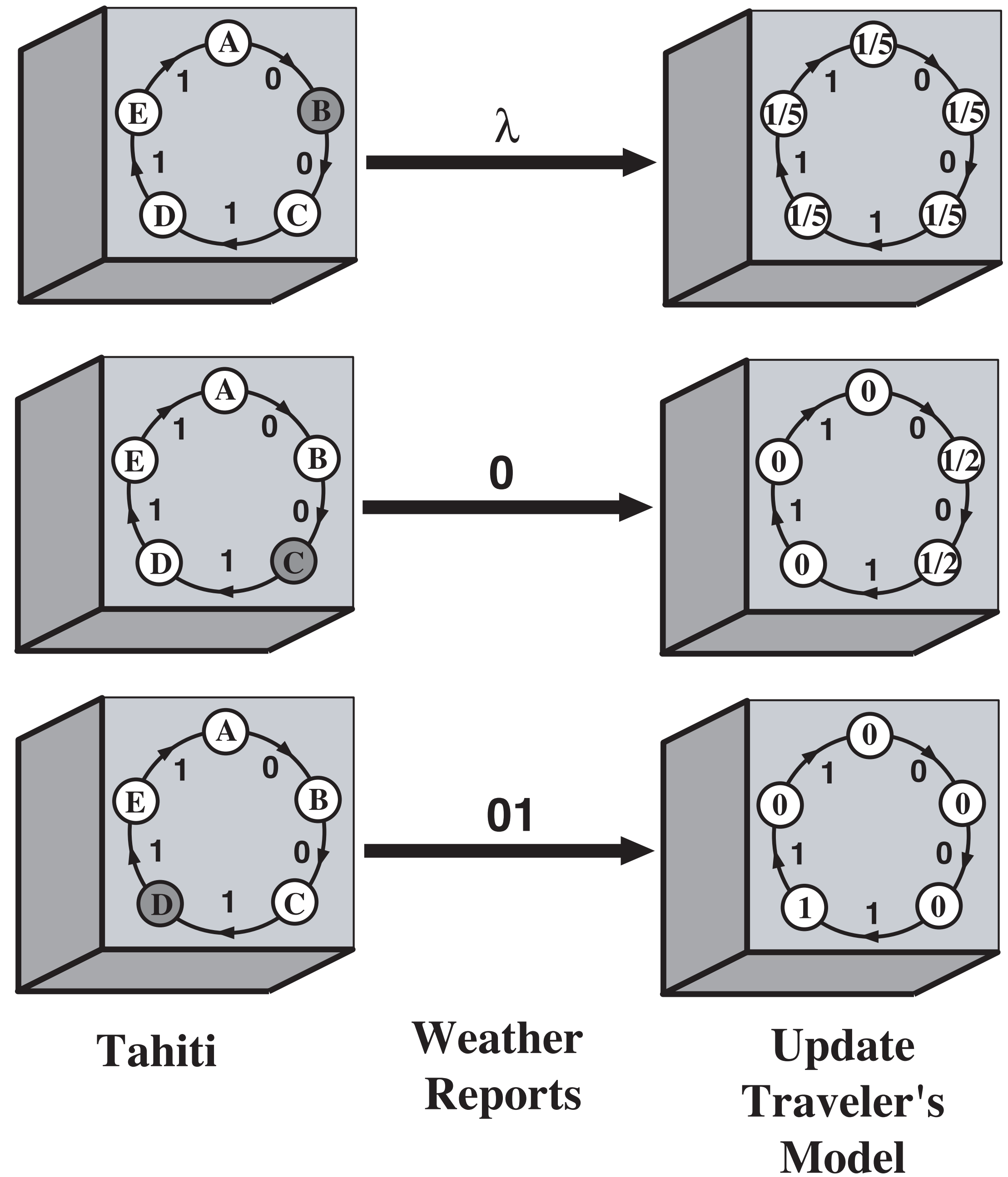
Memory in Processes ...
Example of Transient Information ...
Tahitian Vacation ... packing

0 = Rain
1 = Sun

No weather reports yet.

Rain!

Sun!



Memory in Processes ...

Example of Transient Information ...

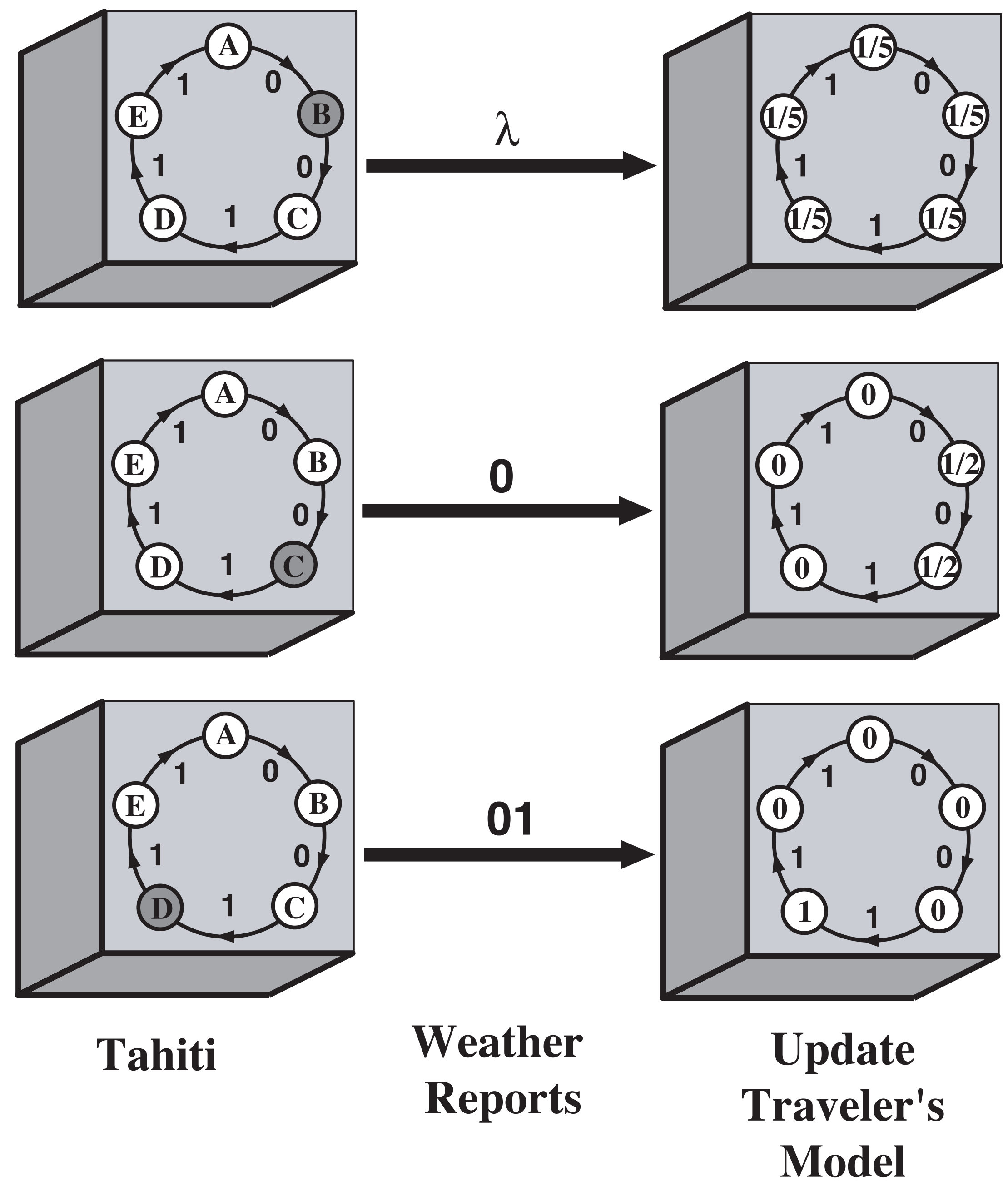
Tahitian Vacation ... packing

0 = Rain
1 = Sun

No weather reports yet.

Rain!

Sun!



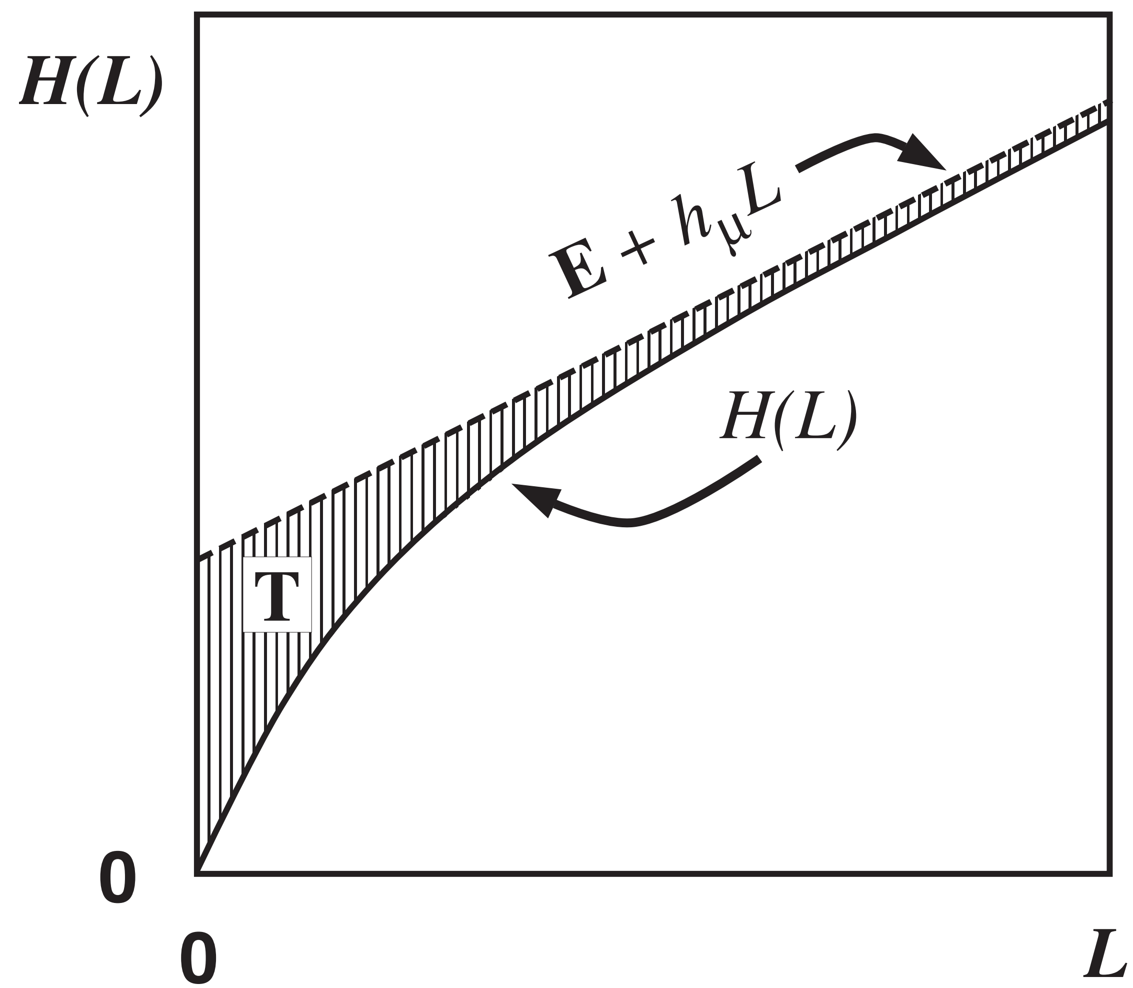
Pack umbrella,
wear shorts on plane

$$\mathbf{T} \approx 4.073 \text{ bit} \times \text{symbols}$$

Memory in Processes ...

Transient Information ...

How to interpret?



Memory in Processes ...

Transient Information ...

Synchronization information:

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

Observer Synchronized to Process:

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

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_μ.

Memory in Processes ...

Transient Information ...

Synchronization information ...

Average state-uncertainty:

$$\mathcal{H}(L) \equiv - \sum_{s^L \in \mathcal{A}^L} \Pr(s^L) \sum_{v \in \mathcal{V}} \Pr(v|s^L) \log_2 \Pr(v|s^L)$$

Synchronization information:

$$\mathbf{S} \equiv \sum_{L=0}^{\infty} \mathcal{H}(L)$$

Memory in Processes ...

Transient Information ...

Synchronization information ...

Theorem: For a R-block Markov process,
the synchronization information is given by:

$$\mathbf{S} = \mathbf{T} + \frac{1}{2} R(R+1) h_\mu$$

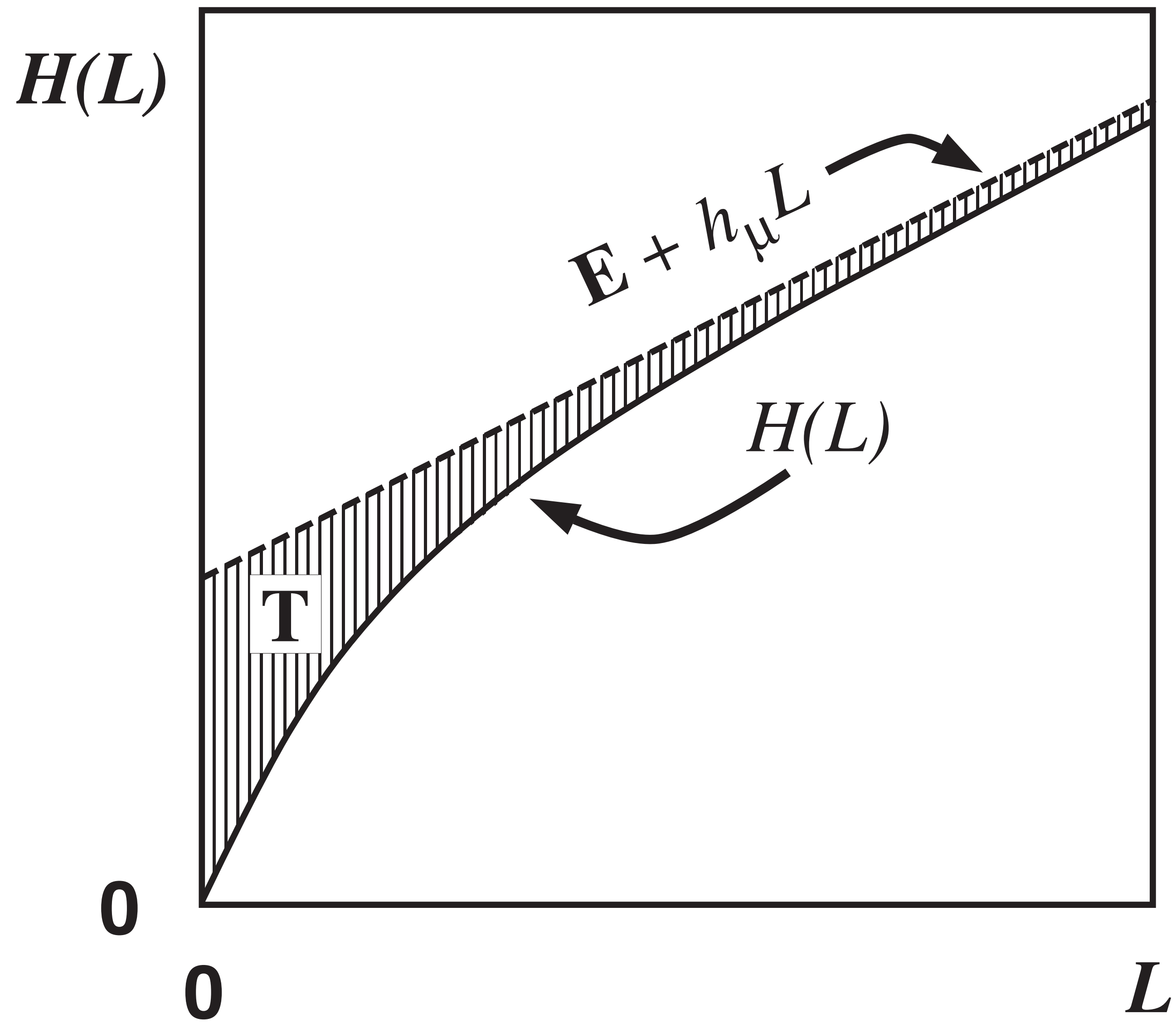
Corollary: For periodic process:

$$\mathbf{S} = \mathbf{T}$$

Memory in Processes ...

Transient Information ...

How to interpret?



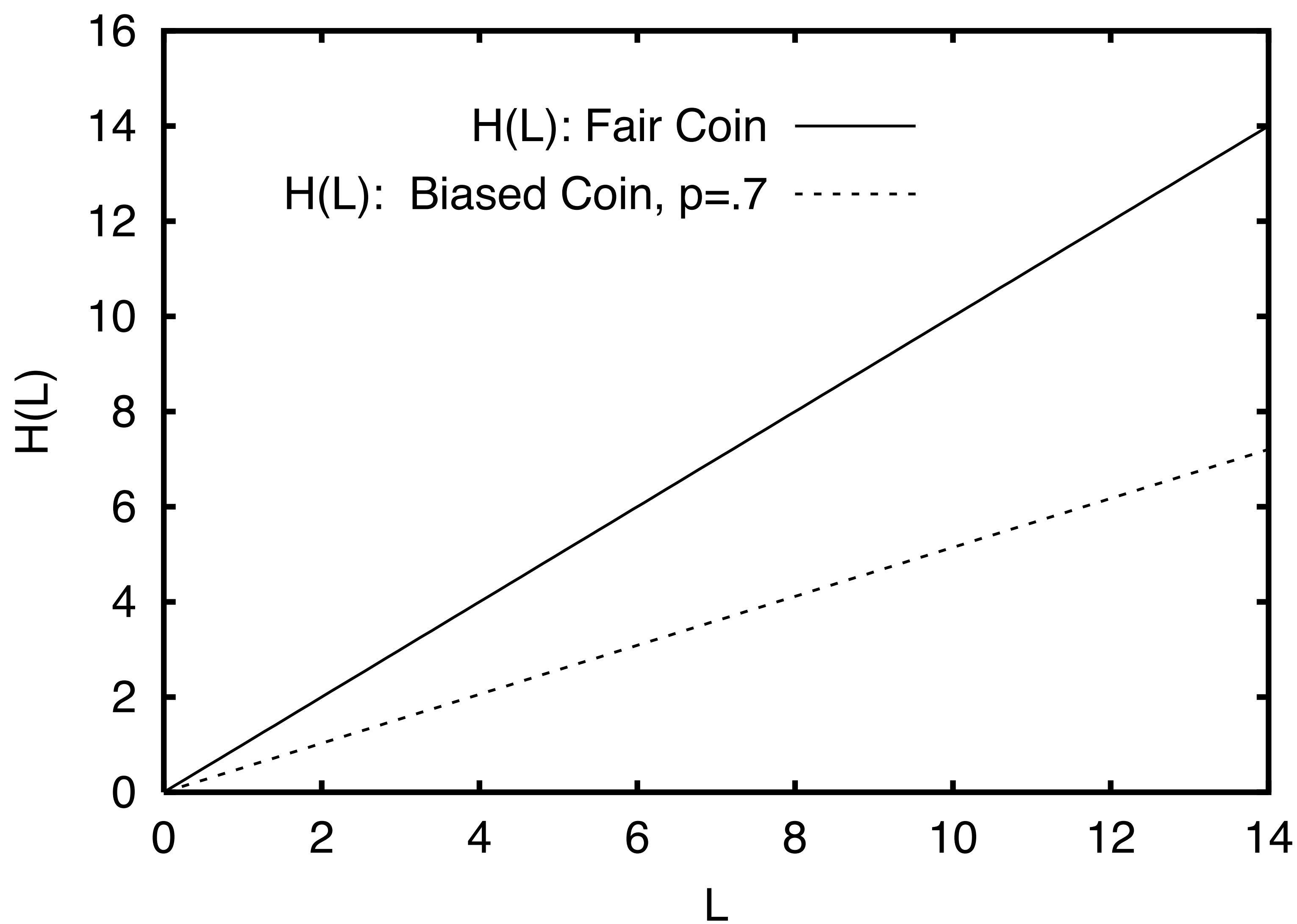
1. Total uncertainty observed while synchronizing.

2. Information to extract to be synchronized.

Memory in Processes ...

Examples of Transient Information:

Fair & Biased Coins
& IID Processes: $\mathbf{T} = 0$



Memory in Processes ...

Examples of Transient Information ...

Period-5 Processes:

There are three distinct:

$$(11000)^\infty$$
$$(10101)^\infty$$
$$(10000)^\infty$$

All:

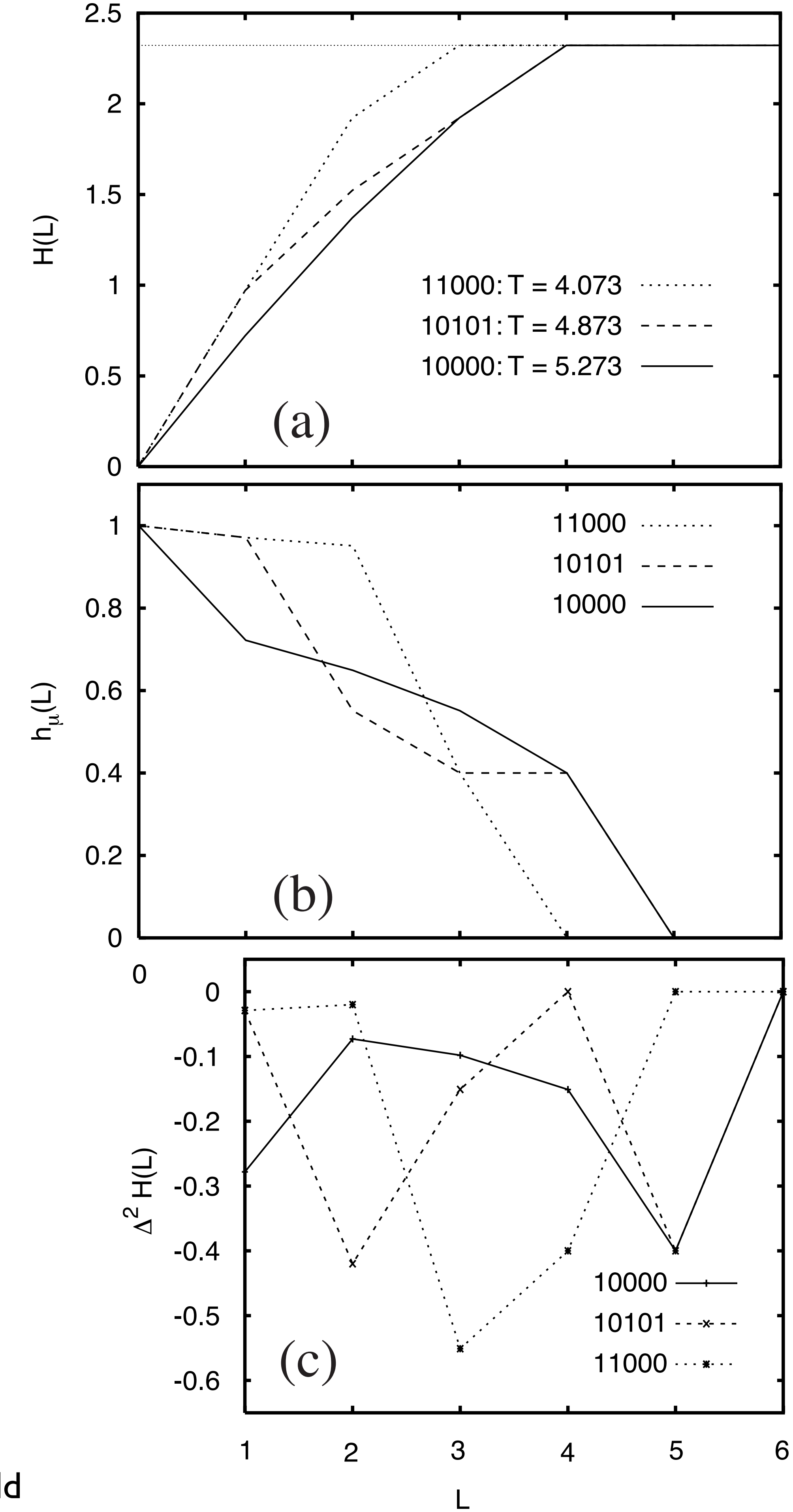
Predictable: $h_\mu = 0 \text{ bits}$

Memory: $\mathbf{E} = \log_2 5 \text{ bits}$

Memory in Processes ...
Examples of Transient Information ...

Period-5 Processes ...

But different ways to sync:



Complexity Lecture 1: Processes and Information theory (CSSS 2011); Jim Crutchfield

Memory in Processes ...

Examples of Transient Information ...

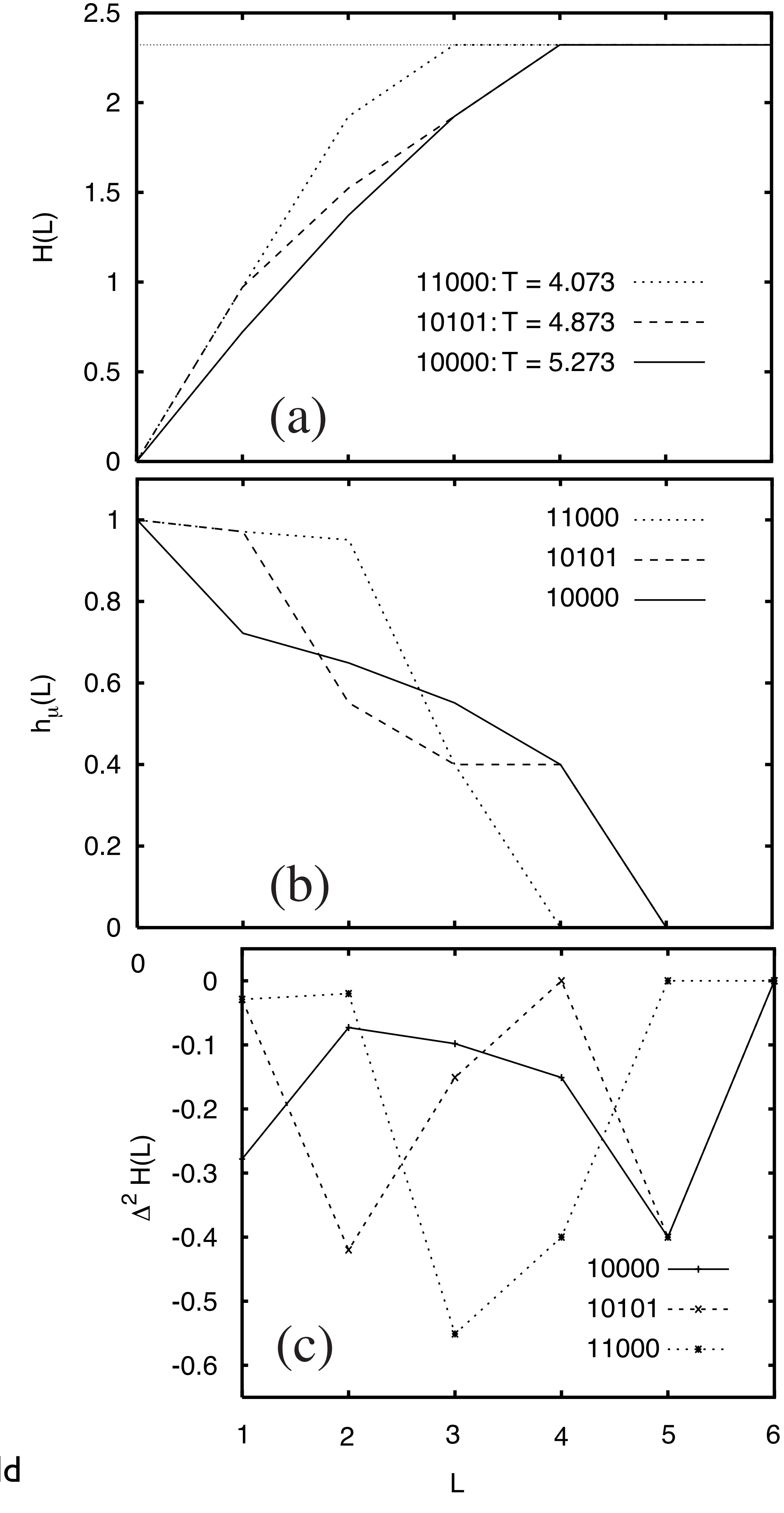
Period-5 Processes ...

But different ways to sync:

$$(11000)^\infty \;\; \mathbf{T} \approx 4.073 \text{ bit} \times \text{symbols}$$

$$(10101)^\infty \;\; \mathbf{T} \approx 4.873 \text{ bit} \times \text{symbols}$$

$$(10000)^\infty \;\; \mathbf{T} \approx 5.273 \text{ bit} \times \text{symbols}$$



Memory in Processes ...

Examples of Transient Information ...

Period-P Processes:

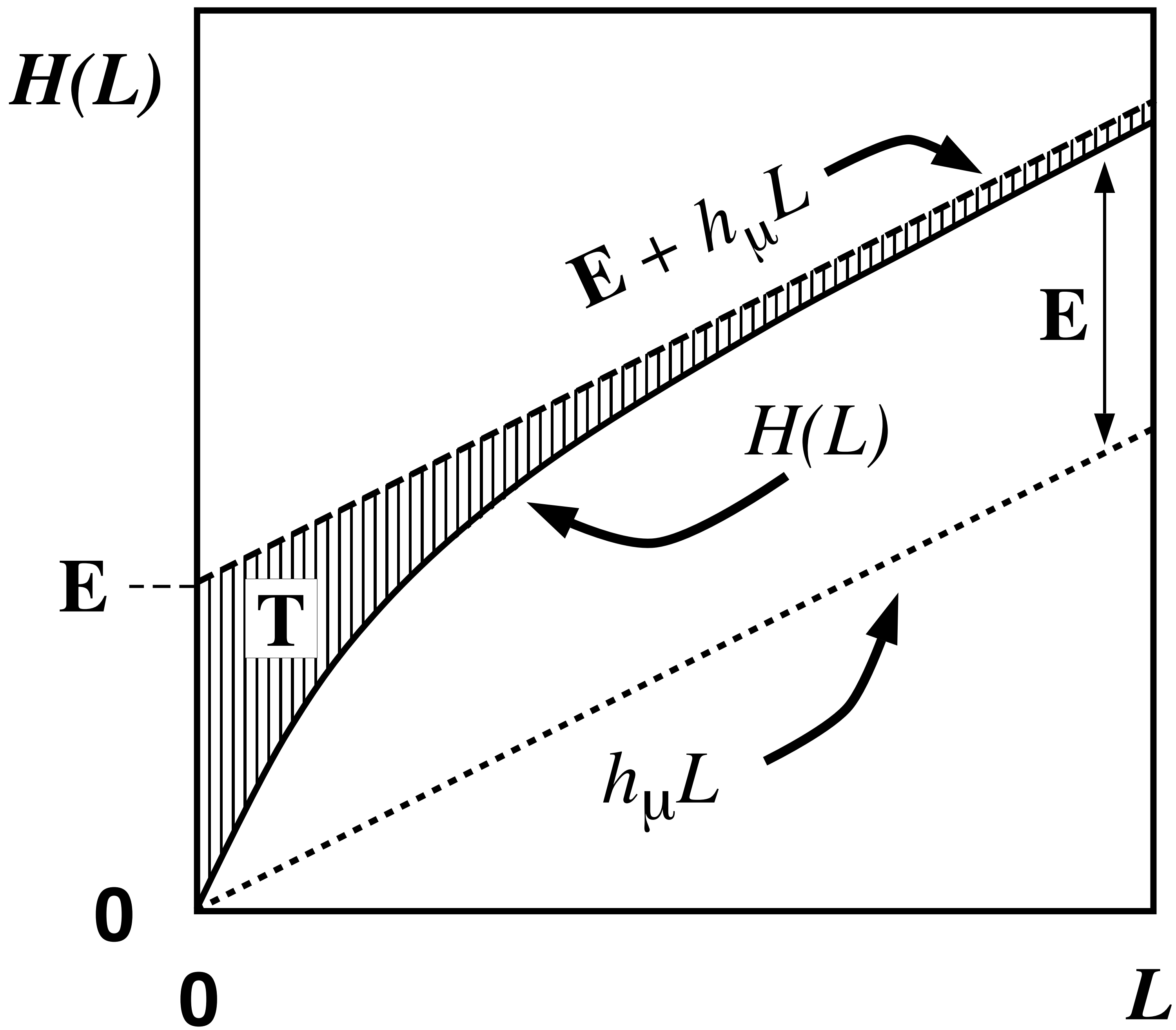
Entropy rate vanishes.

Excess entropy same for all.

But T distinguishes periodic processes.

Memory in Processes ...

Information-Entropy Roadmap for a Stochastic Process:



Complexity Lecture 1: Processes and Information theory (CSSS 2011); Jim Crutchfield

Memory in Processes ...

Regularities Unseen, Randomness Observed:

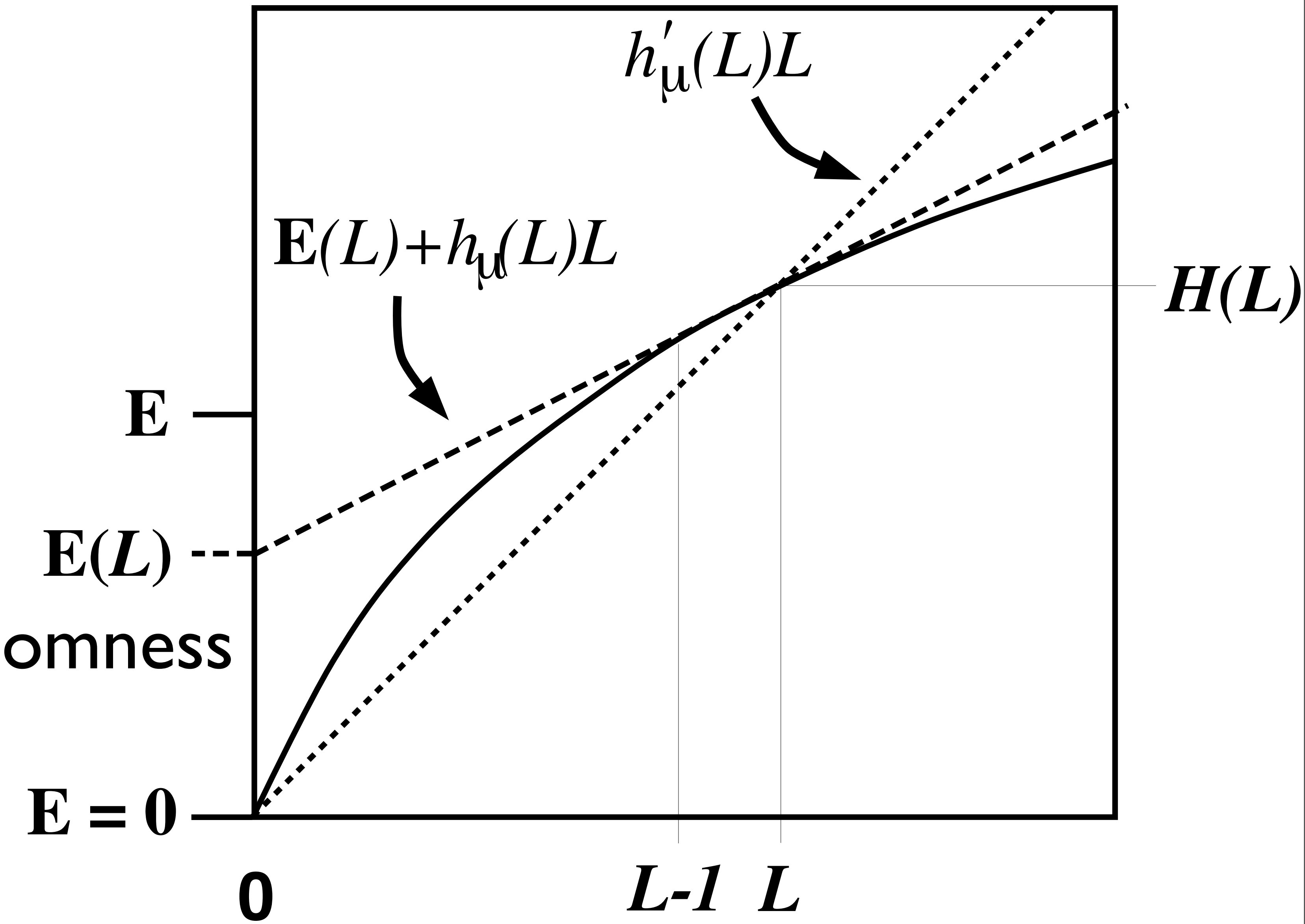
Ignore process's memory

By assuming

$$\mathbf{E} = 0$$

Over-estimate true randomness

$$h_\mu{}' > h_\mu$$



Lesson:

Structure $(\mathbf{E} \ \& \ \mathbf{T})$ converted to apparent randomness (h_μ).

Memory in Processes ...

Calculus of the Entropy Hierarchy:

Via Discrete-Time Derivatives and Integrals

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

Memory in Processes ...

What is information?

Depends on the question!

Uncertainty, surprise, randomness, ...
Compressibility.
Transmission rate.
Memory, apparent stored information, ...
Synchronization.
...

Memory in Processes ...

RGJ Lecture:

Process info diagram

Anatomy of a Bit:
Process i-Diagram
Meaning of the atoms