Beyond the Typical Set: Fluctuation Spectroscopy

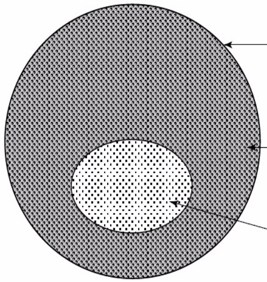
Cina Aghamohammadi

Complexity Sciences Center
Department of Physics
University of California, Davis

Adviser: James P. Crutchfield

June 2, 2015

Typical Set



$$A_\epsilon^n = \{(x_1, x_2, ..., x_n) : 2^{-n(H(X)+\epsilon)} \leq P(x_1, x_2, ..., x_n) \leq 2^{-n(H(X)-\epsilon)}\}$$

For large n, typical set is most probable, and the probability of each sequence in the typical set, A_ϵ^n, have almost the same value $2^{-nH(X)}$.

Information Measures

All the information measures are defined on the typical set!

Info Measures

- $h_\mu = \mathrm{H}[X_0|X_{:0}]$
- $E = \mathrm{I}[X_{:0}; X_{0:}] = \mathrm{I}[\mathcal{S}^-; \mathcal{S}^+]$
- $r_\mu = H[X_0|X_{:0}, X_{1:}]$
- $b_\mu = \mathrm{I}[X_0; X_{1:}|X_{:0}]$

But what about the non typical part?
There are really rare.
Events and their probabilities lying outside typical set are fluctuations or, sometimes, deviations.

Non Typical Sets $\sim$ Rare events

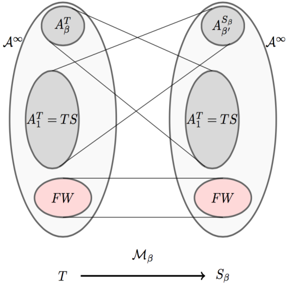
Goal: we want to have information measures for all of the parts of the whole sets.

What is the meaning of that?

For example we know number of words in typical set grows as $\exp(h_\mu L)$ and their probabilities decay as $\exp(-h_\mu L)$.

we can ask a same question for other parts of the whole set.

Idea: β mapping



How to calculate information measures for a subset of $\mathcal{A}^\infty$ (e.g., $\mathcal{A}^\beta$)?
If we could find a mapping that map our process T to new process S_β in a way that it's typical set be $\mathcal{A}^\beta$ then we could calculate all the information measures for S_β and that gives us the answer.

Partitioning the whole set.

How?

because we map $\mathcal{A}^\infty$ to $\mathcal{A}^\beta$ and all the members of $\mathcal{A}^\infty$ have the same decay rate for probability then all the members of the $\mathcal{A}^\beta$ should have the same decay rate too.

so?

We put all the words with same decay rate of probability in a same partition and label that partition with β

Fluctuation spectroscopy

To each word $w \in \mathcal{A}^\ell$ one associates an energy density:

$$U_w^\ell := \frac{-\log_2 \Pr(w)}{\ell} ,$$

mirroring the Boltzmann weight common in statistical physics: $\Pr(w) \propto e^{-U(w)}$.
Naturally, different words w and v may lead to same energy density, $U_w^\ell = U_v^\ell$. And so, in the set $U^\ell = \{U_w^\ell : w \in \mathcal{A}^\ell\}$, energy values may appear repeatedly. Let's denote the frequency of equal U_w^ℓs by $N(U_w^\ell)$. Then, for the thermodynamic macrostate at energy U, we define the *thermodynamic entropy density*:

$$S(U) := \lim_{\ell \to \infty} \frac{\log_2 N(U_w^\ell = U)}{\ell}$$

The Map

The Map

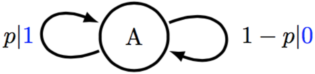
- $\left(T_\beta^{(x)}\right)_{ij} = e^{\beta \ln \Pr(x|\sigma_i)} = \left(\Pr(x|\sigma_i)\right)^\beta$
- $\mathbf{T}_\beta = \sum_{x \in \mathcal{A}} T_\beta^{(x)}$
- $\mathbf{l}_\beta \mathbf{T}_\beta = \lambda_\beta \mathbf{l}_\beta, \mathbf{T}_\beta \mathbf{r}_\beta = \lambda_\beta \mathbf{r}_\beta$
- $\mathbf{l}_\beta \cdot \mathbf{r}_\beta = 1$

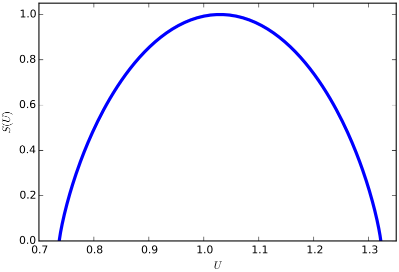
We drove the correct mapping!
$\mathcal{M}_\beta : \mathbf{T} \to \mathbf{S}_\beta$ given by:

$$(\mathbf{S}_\beta)_{ij} = \frac{(\mathbf{T}_\beta)_{ij}(\widehat{\mathbf{r}}_\beta)_j}{\widehat{\lambda}_\beta(\widehat{\mathbf{r}}_\beta)_i} \; ,$$

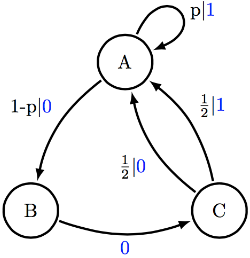
$$\left(\mathbf{S}_\beta^{(x)}\right)_{ij} = \frac{\left(\mathbf{T}_\beta^{(x)}\right)_{ij}(\widehat{\mathbf{r}}_\beta)_j}{\widehat{\lambda}_\beta(\widehat{\mathbf{r}}_\beta)_i} \; .$$

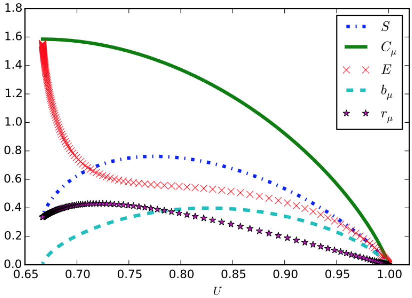
Biased Coin



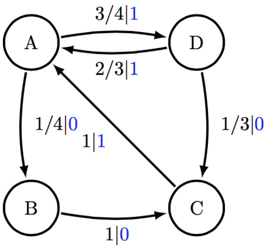


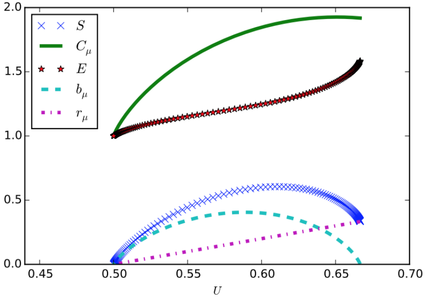
Nemo ~ persistent symmetry



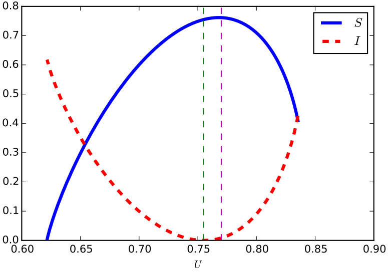


RRIP ∼ hidden symmetry





Large deviation

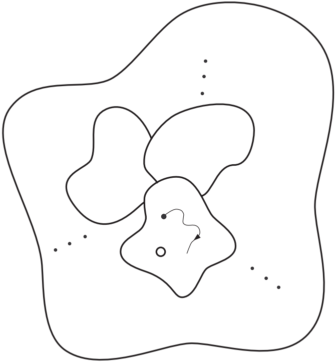


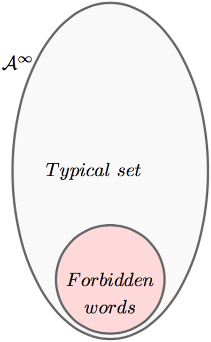
Large deviation rate (How each partition decay?):

$I(U) := \lim_{L\to\infty}\left[-\frac{\log_2 \Pr(U^L)}{L}\right]$

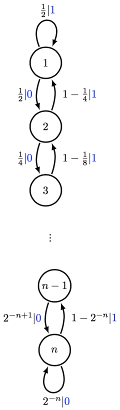
It could be shown that $I(U) = U - S(U)$

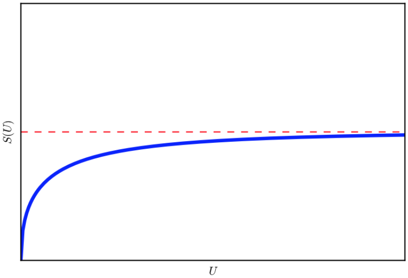
Thermodynamic Classes in Process Space





Infinite-State Processes





Non Ergodicity