Note: More on the Topological Entropy Rate

36-462, Spring 2009

27 January 2009
to accompany lecture 5

The topological entropy rate is a basic measure of how much flexibility there is in the dyamics — how many different kinds of patterns it can produce, and how much the past of the process constrains its future behavior. This note tries to clarify this idea.

In lecture, we considered the symbolic dynamics we get by combining a continuous map Φ with a discrete partition $\mathcal{B}$ — we turn continuous trajectories in the state space, $S_1, S_2, S_3, \ldots$, into discrete symbol sequences, $X_1, X_2, X_3, \ldots \equiv X_1^\infty$; we said that $X_t = b(S_t)$. We also noted that X_1^∞ is completely determined by the initial state, S_1. We will say $X_1^\infty = b_1^\infty(S_1)$.

A **word** is a finite sequence of symbols. A word is **allowed** by the symbolic dynamics if it can occur as a subsequence. That is, a word w of length n is allowed if there is some s and some t such that $b_t^{t+n-1}(s) = w$. If a word is not allowed then it is **forbidden**.

Every allowed word w of length n has a "parent" which is a word of length $n-1$ — this is the word formed of its first $n-1$ symbols. So the number of allowed words of length n must be at least equal to the number of allowed words of length $n-1$. On the other hand there could be more than one way to continue a given word, i.e., two or more words of length n could share the same **prefix** of $n-1$ symbols. Thus at least sometimes the number of allowed words grows with n. In fact, for many symbolic dynamical systems, the number of allowed words grows exponentially with n.

Let's write W_n for the number of allowed words of length n. In lecture we defined the **topological entropy rate** to be[1]

$$h_0 \equiv \lim_{n\to\infty} \frac{1}{n} \log W_n \tag{1}$$

(If you want to know why this is written h_0, see the notes by Feldman on information theory, in the readings.) This is, as we said, the long-run rate of exponential growth in the number of allowed words. Let's first re-write Eq. 1 in a way which may make that more transparent, and then look at its values for different kinds of dynamics.

[1] Writing the subscript on $\log_2$ all the time gets tiresome, so all logarithms will be to base 2.

Begin by doing something pointless-looking, which is to add and subtract a lot of canceling terms, creating a "telescoping sum":

$$\begin{aligned}
\log W_n &= && (2)\\
&= \log W_n - \log W_{n-1} + \log W_{n-1} - \log W_{n-2} + \log W_{n-2} - \ldots - \log W_1 + \log W_1 - \log 1 \\
&= \sum_{t=1}^{n} \log W_t - \log W_{t-1} && (3)\\
&= \sum_{t=1}^{n} \log \frac{W_t}{W_{t-1}} && (4)
\end{aligned}$$

where we conveniently define $W_0 = 1$, and remember of course that $\log 1 = 0$. Now stick Eq. 4 into 1.

$$h_0 = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \log \frac{W_t}{W_{t-1}} \tag{5}$$

$\frac{W_t}{W_{t-1}}$ is the ratio by which the number of allowed words expands when going from $t-1$ to t. So this looks rather like the time-average of the log of that expansion factor. This suggests that h_0 tells us by what factor, on average, the number of allowed words keeps expanding. This is why we can, roughly, think of 2^{h_0} as the number of choices we have for how to extend a word which is already very long. — Note however that h_0 is not a *real* average over time, since W_t isn't something we calculate for a *particular* state, but rather is a property of the whole state *space*.

One place we have already seen a (real) time-average of a log expansion factor is the Lyapunov exponent in one dimension:

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \log |\Phi'(x_t)| \tag{6}$$

This might lead you to guess that there is some kind of connection between λ and h_0. This guess is sometimes right (which is why I bring it up, of course). In one dimension, it turns out that $h_0 \geq \lambda$. This is not obvious, and a proper demonstration is pretty intricate.

How does this play out in terms of the kind of dynamics we already know about?

Attracting fixed point Any trajectory *at* a fixed point must stay in the same cell of the partition forever, so a fixed point s^* produces a symbol sequence which is an infinite repetition of a single letter. Since all trajectories are attracted to the fixed point, eventually they all enter the same cell of the partition as the fixed point and stop changing. So, while W_n never shrinks, eventually $W_n = W_{n+1}$, and so from Eq. 5 we can see that h_0 must be zero.

Attracting limit cycle Cycles can produce more than one word because we can start them in different phases. A cycle of period p can produce, at most, p distinct words of any length. ("At most", because multiple points in the cycle can belong to the same cell of the partition. For instance, with a non-generating partition, *every* point of the cycle could be in the same cell.[2]) Now the same kind of argument applies as in the previous case — having p different allowed words instead of just one doesn't really change things. The important point is that after a certain length, the number of allowed words isn't growing.

Chaotic attractor What we want to say is that $h_0 > 0$ on a chaotic attractor. This requires a little care. If the partition $\mathcal{B}$ is chosen stupidly, it could put the *entire* attractor in a single cell. Then $h_0 = 0$. So let's suppose the attractor occupies at least two cells. What we want to say is that no matter how close s and r might be, eventually, for some n, $\Phi^n(s)$ and $\Phi^n(r)$ will be on totally different places on the attractor, in particular one of them will be in one cell and the other will be in the other cell. So even if $b_1^{n-1}(s) = b_1^{n-1}(r)$, we have $b_n(s) \neq b_n(r)$. This means there is at least one word of length $n - 1$ which can be extend into at least two words of length n. So $W_n/W_{n-1} > 1$. Similarly, there will be another point in the state space, call it q, which is closer to s and so $b_1^n(s) = b_1^n(q)$, but eventually the orbit of q will diverge from the orbit of s, so at some time $n + m$, $b_{n+m}(s) \neq b_{n+m}(q)$, meaning that $W_{n+m}/W_n > 1$. And so on.

While plausible, this line of argument is only roughly right — to really turn it into a proof one would need to nail down a lot of details, like showing that nearby trajectories always have to *eventually* wind up on opposite sides of a cell boundary. But the spirit of it is right, for not-too-unfortunate partitions.

To sum up: periodic attractors mean $h_0 = 0$, because eventually there is no choice about how to continue any symbol sequence. Chaotic attractors (pretty much) need $h_0 > 0$, because sensitive dependence on initial conditions will eventually give us a choice between symbols, and once it does that once it keeps doing so over and over, so even asymptotically there is always still some room for choices, and the number of allowed words grows exponentially.

There are weird situations where $h_0 = 0$ but the behavior is not periodic, because W_n grows forever with n, but less than exponentially — say $W_n \propto n^2$. You will not, inshallah, have to deal with such things.

[2] Convince yourself that this is not possible with a generating partition.