Summer 2012 NSERC USRA Report

Entropy in Symbolic Dynamics

John Enns

This summer I had the pleasure of working with Professors Brian Marcus and Andrew Rechnitzer of the UBC Math Department along with my partner Nigel Burke in the area of symbolic dynamics. The fundamental object which we studied was the entropy of a shift. Given a directed edge-labeled graph $G = (V, E)$ we say that G presents the (1-dimensional) constrained system X, which is the set of bi-infinite sequences of labels which can be read off along paths of G. For $n \in \mathbb{N}$ we also write $B_n(X)$ for the set of such sequences of length n. This notion of constrained system corresponds to the idea of a sofic shift in symbolic dynamics. For example, the directed graph shown in Figure 1 below presents the constrained code $X = \{x \in \{0,1\}^{\mathbb{Z}} : x \text{ contains no consecutive 1s}\}$, which is known as the golden mean shift.

The entropy of a constrained code or shift is defined as

$$h(X) = \lim_{n\to\infty} \frac{\log_2 |B_n(X)|}{n}. \tag{1}$$

The entropy of a shift corresponds roughly to how many choices of symbols one has per site in some average sense. The problem of calculating the entropy of a 1D shift is completely solved. Given a graph presentation G of X, $h(X)$ is equal to the log of the largest eigenvalue of the adjacency matrix of G. For example, using the graph in Figure 1, the entropy of the golden mean shift is quickly calculated to be $\log \frac{1+\sqrt{5}}{2}$. This theory is developed in [1] and [2], for example.

On the other hand, the corresponding problem of calculating the entropy in dimensions higher than 1 is completely unsolved, except for a few specific cases. For example, if X is the golden mean shift, write $B_{k,n}(X \otimes X)$ for the set of $k \times n$ grids of 0s and 1s such that no 1s are vertically or horizontally adjacent. The value of the limit

$$h(X \otimes X) = \lim_{k,n\to\infty} \frac{\log_2 |B_{k,n}(X \otimes X)|}{kn}, \tag{2}$$

which represents the 2D entropy, is unknown. Appropriately, techniques have been developed to find rigorous bounds or approximations to higher dimensional entropies, and we studied some of these this summer.

The limit in (2) is independent of how $k, n \to \infty$, and in fact one may take iterated limits. Thus the entropy may be approximated by computing the entropy $\log_2 \Lambda_k$ for a strip of height k via 1D techniques, and using the fact that $h(X \otimes X) = \lim_{k\to\infty} \frac{\log_2 \Lambda_k}{k}$. The approximations $\frac{\log_2 \Lambda_k}{k}$ that one obtains to the entropy using this method are not good, and storing the matrices required to compute the eigenvalues Λ_k quickly becomes taxing. Instead, the most well-known techniques for obtaining rigorous bounds on the 2D entropy use clever modifications of these ideas (see [3]). These techniques can also be applied in dimensions higher than 2, as is done in [4].

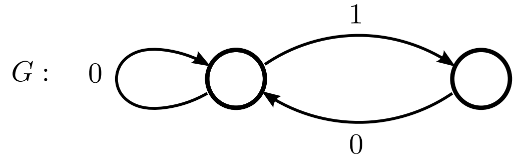

Figure 1: A presentation G of the golden mean shift.

We studied variants of these techniques in several papers and implemented them in code in different scenarios. In doing so, we noticed a novel method of obtaining approximations to 2D entropies - namely, if one computes a sequence of eigenvalues not for configurations of symbols on strips, but for configurations of symbols on closely related helices, then this sequence of eigenvalues $(\lambda_k)_{k\geq 1}$ appears to tend towards the value of the 2D entropy. We then formalized this as 'the helix method' and proved convergence of the helix eigenvalues to the 2D (or 3D) entropy for a large class of shifts, along with some relatively weak bounds on the entropy based on these eigenvalues. These eigenvalues appeared to be a rather fast converging sequence. What is more, for the 2D golden mean shift problem, we observed the pattern

$$\lambda_{2k} \leq h(X \otimes X) \leq \lambda_{2k+1} \tag{3}$$

for small values of k (as far as we could compute). We conjectured that this pattern continues forever, thus providing us with a sequence of very good bounds.

Unfortunately, it was not meant to be. We discovered a paper in which our helix method is introduced and discussed as 'the 1-vertex transfer method' ([5]). In this paper, the authors arrive at the same conjecture as us. However, later numerical evidence gathered by us for various shifts indicates that the conjecture is probably very dependent on properties specific to the golden mean shift. Thus our conclusion is that the helix method is most likely only useful as a way of obtaining approximations to entropies, rather than bounds.

We continued on by studying some simple higher-dimensional shifts for which the entropy is known exactly. One which caught our eye was the odd shift, Θ. This shift is defined as

$$\Theta = \{x \in \{0,1\}^{\mathbb{Z}} : \text{ there are an odd number of 0s between consecutive 1s in } x\}. \tag{4}$$

A combinatorial argument shows that the entropy of Θ is equal to $1/2$ in any number of dimensions. However, its counterpart, the even shift (defined similarly), has unknown entropy in any dimension larger than 1. It is well known that the entropy cannot increase as one moves up a dimension. This lead us to ask the question What kind of shifts X have the property that $h(X \otimes X) = h(X)$? We managed to characterize all shifts which have this property, with some slight assumptions. Without going into details, the rough answer is that X must not place any restrictions on allowed sequences of symbols, except for a kind of periodicity. A consequence of this result is that if $h(X \otimes X) = h(X)$ then the entropy must be the same in any number of dimensions.

The answer to the question of whether there exists such a characterization in higher dimensions is no. We found counterexamples to the direct generalization of this theorem. Of course, that does not preclude the possibility of a more complicated statement in higher dimensions, but that question, like many in the study of higher dimensional shifts, remains open.

I had an excellent time this summer and I learned a lot about entropy and symbolic dynamics in general. I am going to continue to work with my supervisors during the fall term. We are planning on writing some of these results into a paper.

References

[1] B. H. Marcus, R. R. Roth and P. H. Siegel, *An Introduction to Coding for Constrained Systems*, Online course notes available from http://www.math.ubc.ca/~marcus/Handbook/index.html

[2] D. Lind and B. H. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.

[3] N. J. Calkin and H. S. Wilf, *The number of independent sets in a grid graph*, SIAM J. Discrete Math. **11** (1998) 54 - 60.

[4] Z. Nagy and K. Zeger, *Capacity bounds for the 3-dimensional (0,1) run length limited channel*, IEEE Trans. Info. Theory. **46** (2000) 1030 - 1033.

[5] S. Friedland, P. H. Lundow and K. Markström, *The 1-vertex transfer matrix and accurate estimation of channel capacity*, IEEE Trans. Info. Theory. **56** (2010) 3692 - 3699.