The Entropy of Markov Trajectories

Laura Ekroot, *Member, IEEE*, and Thomas M. Cover, *Fellow, IEEE*

***Abstract*—The idea of thermodynamic depth put forth by Lloyd and Pagels requires the computation of the entropy of Markov trajectories. Toward this end we consider an irreducible finite state Markov chain with transition matrix P and associated entropy rate $H(\mathcal{X}) = -\sum_{i,j}\mu_i P_{ij}\log P_{ij}$, where μ is the stationary distribution given by the solution of $\mu = \mu P$. A trajectory T_{ij} of the Markov chain is a path with initial state i, final state j, and no intervening states equal to j. We show that the entropy $H(T_{ii})$ of the random trajectory originating and terminating in state i is given by $H(T_{ii}) = H(\mathcal{X})/\mu_i$. Thus the entropy of the random trajectory T_{ii} is the product of the expected number of steps $1/\mu_i$ to return to state i and the entropy rate $H(\mathcal{X})$ per step for the stationary Markov chain. A general closed form solution for the entropies $H(T_{ij})$ is given by**

$$H = K - \tilde{K} + H_\Delta,$$

where H is the matrix of trajectory entropies $H_{ij} = H(T_{ij})$; $K = (I - P + A)^{-1}(H^* - H_\Delta)$; $\tilde{K}$ is a matrix in which the ijth element $\tilde{K}_{ij}$ equals the diagonal element K_{jj} of K; A is the matrix of stationary probabilities with entries $A_{ij} = \mu_j$; H^* is the matrix of single-step entropies with entries $H^*_{ij} = H(P_{i\cdot}) = -\sum_k P_{ik}\log P_{ik}$; and H_Δ is a diagonal matrix with entries $(H_\Delta)_{ii} = H(\mathcal{X})/\mu_i$.

I. Introduction

The number of bits of randomness in a trajectory of a Markov chain has applications in backgammon, gambling, population growth, and evolution. Indeed, Lloyd and Pagels [1], [2] (and to some extent Bennett [3]) define notions of logical depth and thermodynamic depth which amount to entropy measures of the process (or path) by which a state arises. For example, the thermodynamic depth of a cat is large, but the thermodynamic depth of a cat and her kitten is not much larger. Here the initial state could be a state of primordial ooze and the final state the cat and her kitten. We are, therefore, interested in the entropy of how the process got to its final state, i.e., the descriptive complexity of the path.

We shall investigate the entropy of trajectories of finite state irreducible Markov chains. Consider a finite state irreducible Markov chain with transition matrix P and initial state $X_1 = i$. The entropy rate

$$H(\mathcal{X}) = \lim_{n\to\infty}\frac{H(X_1, X_2, \cdots, X_n)}{n} \tag{1}$$

is always well defined, since the limit exists, and is given by

$$H(\mathcal{X}) = -\sum_{i,j}\mu_i P_{ij}\log P_{ij}, \tag{2}$$

where μ is the (unique) solution of the equations

$$\mu_j = \sum_i \mu_i P_{ij}, \qquad \text{for all } j. \tag{3}$$

Manuscript received July 24, 1991; revised November 13, 1992. This work was supported in part by the National Science Foundation under NSF Grant NCR-89-14538, by DARPA Contract J-FBI-91-218, and under JSEP Contract DAAL03–91–C-0010. This work was done at Stanford University, Stanford, CA.

L. Ekroot is with the Communications Systems Research Group, Jet Propulsion Laboratory, California Institute of Technology, 238–420, 4800 Oak Grove, Pasadena, CA 91109.

T. Cover is with the Departments of Electrical Engineering and Statistics, Durand 121, Stanford University, Stanford, CA 94305.

IEEE Log Number 9209455.

Definition: A *trajectory* t_{ij} from state i to state j of a Markov chain is a path with initial state i, final state j, and no intervening state equal to j.

The probability $p(t_{ij})$ of a trajectory $t_{ij} = ix_2x_3\cdots x_kj$ is given by

$$p(t_{ij}) = P_{ix_2}P_{x_2x_3}\cdots P_{x_kj}. \tag{4}$$

This is the conditional probability of the trajectory t_{ij} from i to j given $X_1 = i$.

Irreducibility of P implies that

$$\sum_{t_{ij}\in T_{ij}} p(t_{ij}) = 1, \tag{5}$$

where T_{ij} is the set of all trajectories from i to j. Thus, given initial state $X_1 = i$, the random trajectory T_{ij} is a finite length sequence drawn according to the probability mass function $p(t_{ij})$.

Definition: The *entropy* H_{ij} of the trajectory from i to j is defined by

$$H_{ij} = H(T_{ij}) = -\sum_{t_{ij}\in T_{ij}} p(t_{ij})\log p(t_{ij}). \tag{6}$$

Our objective is to determine a closed form expression for H_{ij}.

II. Recurrence Relation

We develop a recurrence relation for H_{ij}. From this we determine H_{ii}, the entropy of the path that returns to the initial state. Subsequently, this solution for H_{ii} will be substituted into the recurrence to complete the determination of H_{ij}, for all i, j.

One can solve for H_{ij} by modifying the transition matrix P to make j an absorbing state. This yields a different recurrence relation for j. Instead, for unity and symmetry, we resolve for the matrix $[H_{ij}]$ directly in terms of P.

Let $P_{i\cdot}$ denote the ith row of the Markov transition matrix P. Thus the entropy of the first step of a trajectory originating in state i is given by

$$H(P_{i\cdot}) = -\sum_j P_{ij}\log P_{ij}. \tag{7}$$

The fundamental recurrence then becomes

$$H_{ij} = H(P_{i\cdot}) + \sum_{k\neq j} P_{ik}H_{kj}, \tag{8}$$

which follows from the chain rule for entropy. Specifically, the entropy of a trajectory is given by the entropy of the first step plus the conditional entropy of the remaining trajectory given the first step.

Alternatively, Theorem 1 can be restated as the matrix recurrence

$$H = H^* + PH - PH_\Delta, \tag{9}$$

where the matrix of trajectory entropies is

$$H = [H_{ij}], \tag{10}$$

the matrix of first step entropies is

$$H^* = \begin{pmatrix} H(P_{1\cdot}) & H(P_{1\cdot}) & \cdots & H(P_{1\cdot}) \\ H(P_{2\cdot}) & H(P_{2\cdot}) & \cdots & H(P_{2\cdot}) \\ \vdots & \vdots & \vdots & \vdots \\ H(P_{m\cdot}) & H(P_{m\cdot}) & \cdots & H(P_{m\cdot}) \end{pmatrix}, \tag{11}$$

and the diagonal matrix associated with H is

$$H_\Delta = \begin{pmatrix} H_{11} & 0 & 0 & \cdots & 0 \\ 0 & H_{22} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & H_{mm} \end{pmatrix}. \quad (12)$$

Theorem 1: For an irreducible Markov chain, the entropy H_{ii} of the random trajectory from state i back to state i is given by

$$H_{ii} = \frac{H(\mathcal{X})}{\mu_i}, \quad (13)$$

where μ_i is the stationary probability for state i and $H(\mathcal{X})$ is the entropy rate given in (2).

Proof: The matrix recurrence is

$$H = H^* + PH - PH_\Delta. \quad (14)$$

Multiplying on the left by the stationary distribution μ yields

$$\mu H = \mu H^* + \mu PH - \mu PH_\Delta. \quad (15)$$

Applying $\mu P = \mu$, it follows that

$$\mu H = \mu H^* + \mu H - \mu H_\Delta. \quad (16)$$

Thus,

$$\mu H^* = \mu H_\Delta. \quad (17)$$

But μH^* is a vector with equal components each given by $\sum_k \mu_k H(P_{k\cdot}) = H(\mathcal{X})$, where $H(\mathcal{X}) = -\sum_{k,j} \mu_k P_{kj} \log P_{kj}$ is the entropy rate of the Markov chain given in (2). Equation (17) may then be written equivalently as

$$(H(\mathcal{X}), H(\mathcal{X}), \cdots, H(\mathcal{X})) = (\mu_1 H_{11}, \mu_2 H_{22}, \cdots, \mu_m H_{mm}). \quad (18)$$

By equating components, it follows that

$$H_{ii} = \frac{H(\mathcal{X})}{\mu_i}. \quad (19)$$

□

The entropy of the trajectory T_{ii} has an interpretation as the product of the expected number of steps $1/\mu_i$ to return to state i and the per-step entropy rate $H(\mathcal{X})$ of the stationary Markov chain. Parsing a sequence of states generated by the Markov chain with initial state i into trajectories $T_{ii}^{(1)} T_{ii}^{(2)} \cdots$ and applying a law of large numbers argument makes it obvious that the entropy rate of the $\{T_{ii}^{(k)}\}_{k=1}^{\infty}$ process must equal $H(\mathcal{X})/\mu_i$. A similar result for randomly stopped sequences is found in [4]. However, the example in Section V can be used to show that the entropy $H(T_{ij})$ of a trajectory from state i to state j is generally not equal to $H(\mathcal{X})M_{ij}$, where M_{ij} is the expected length of the trajectory from i to j.

III. Example

Consider the irreducible Markov chain depicted in Fig. 1 with transition matrix

$$P = \begin{pmatrix} 0 & 0.9 & 0 & 0.1 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (20)$$

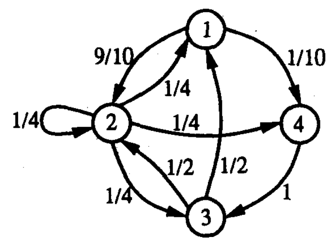

Fig. 1. State transition diagram.

This chain has a unique stationary distribution $\mu = (0.22, 0.42, 0.23, 0.13)$ and an entropy rate of

$$H(\mathcal{X}) = -\sum_{i=1}^{4} \mu_i \sum_{j=1}^{4} P_{ij} \log_2 P_{ij} \quad (21)$$

$$= 1.18 \text{ bits per transition.} \quad (22)$$

The matrix of first step entropies is

$$H^* = \begin{pmatrix} 0.46 & 0.46 & 0.46 & 0.46 \\ 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (23)$$

For example, the entropy of the first step leaving state 1 is $H(0.1, 0.9) = 0.46$ bits. By Theorem 1, the matrix H_Δ is

$$H_\Delta = \begin{pmatrix} 5.32 & 0 & 0 & 0 \\ 0 & 2.80 & 0 & 0 \\ 0 & 0 & 5.07 & 0 \\ 0 & 0 & 0 & 9.25 \end{pmatrix}. \quad (24)$$

Thus, for example, the entropy of the random trajectory from state 4 back to state 4 is 9.25 bits.

IV. General Solution

We now solve the recurrence relation

$$H = H^* + PH - PH_\Delta \quad (25)$$

using the previously determined diagonal values

$$H_{ii} = \frac{H(\mathcal{X})}{\mu_i} \quad (26)$$

to derive the following general theorem.

Theorem 2: If P is the transition matrix of an irreducible finite state Markov chain, then the matrix H of trajectory entropies is given by

$$H = K - \tilde{K} + H_\Delta, \quad (27)$$

where

$$K = (I - P + A)^{-1}(H^* - H_\Delta), \quad (28)$$

$$\tilde{K}_{ij} = K_{jj} \quad \text{for all } i, j, \quad (29)$$

$$A_{ij} = \mu_j \quad \text{for all } i, j, \quad (30)$$

$$H_{ij}^* = H(P_{i\cdot}) \quad \text{for all } i, j, \quad (31)$$

$$\text{and} \quad (H_\Delta)_{ij} = \begin{cases} H(\mathcal{X})/\mu_i, & i = j \\ 0, & i \neq j. \end{cases} \quad (32)$$

Proof: The proof will be separated into three parts. We first derive a solution to the recurrence relation in (25) for aperiodic Markov chains, then extend the proof to periodic Markov chains, and finally prove the uniqueness of the solution.

Rewriting (25), we have

$$H - H_\Delta = (H^* - H_\Delta) + P(H - H_\Delta) \quad (33)$$

or

$$\hat{H} = (H^* - H_\Delta) + P\hat{H} \quad (34)$$

where

$$\hat{H} = H - H_\Delta. \quad (35)$$

Here $\hat{H}$ is the matrix of off-diagonal terms of the matrix of trajectory entropies H. Note that H^* is the known matrix of first step entropies, H_Δ is given by Theorem 1 and $\hat{H}$ has zeros on the diagonal. The fact that $\hat{H}$ has zeros on the diagonal means that there are more equations than unknowns and thus the obvious matrix inverse technique for solving this equation will not work.

However, iteratively substituting for $\hat{H}$ in the recurrence (34), we have

$$\hat{H} = P\hat{H} + (H^* - H_\Delta) \quad (36)$$
$$= P^2\hat{H} + (H^* - H_\Delta) + P(H^* - H_\Delta) \quad (37)$$
$$\vdots \quad (38)$$
$$= P^n\hat{H} + (I + P + P^2 + \cdots + P^{n-1})(H^* - H_\Delta). \quad (39)$$

As $I+P+P^2+\cdots+P^{n-1}$ does not converge, we use the relationship $A(H^* - H_\Delta) = 0$, as proved in (17) of Theorem 1, to write (39) in the equivalent form

$$\hat{H} = P^n\hat{H} + (I + P - A + P^2 - A + \cdots + P^{n-1} - A) \cdot (H^* - H_\Delta). \quad (40)$$

Note that

$$P^n - A = (P - A)^n, \quad (41)$$

which can be proved by applying the fact that $AP = PA = A$ to a binomial expansion of $(P - A)^n$. Substituting (41) into (40), we obtain

$$\hat{H} = P^n\hat{H} + \sum_{k=0}^{n-1}(P - A)^k(H^* - H_\Delta). \quad (42)$$

To proceed, we first assume P is aperiodic, in which case $P^n \to A$. Taking the limit in (42) yields

$$\hat{H} = A\hat{H} + (I - (P - A))^{-1}(H^* - H_\Delta), \quad (43)$$

or equivalently

$$(I - A)\hat{H} = (I - (P - A))^{-1}(H^* - H_\Delta). \quad (44)$$

Note that $(I - A)$ is not invertible. We now denote the right-hand side in (44) by

$$K = (I - P + A)^{-1}(H^* - H_\Delta), \quad (45)$$

and solve

$$(I - A)\hat{H} = K \quad (46)$$

entry by entry using the fact that $\hat{H}_{ii} = 0$.

The ijth component of (46) is

$$\hat{H}_{ij} - \sum_r A_{ir}\hat{H}_{rj} = K_{ij}. \quad (47)$$

Recalling $A_{ir} = \mu_r$, we have

$$\hat{H}_{ij} - \sum_r \mu_r\hat{H}_{rj} = K_{ij}. \quad (48)$$

Since $\hat{H}_{jj} = 0$, the expression for the jjth component is

$$-\sum_r \mu_r\hat{H}_{rj} = K_{jj}. \quad (49)$$

Substituting (49) back into (48) yields

$$\hat{H}_{ij} = K_{ij} - K_{jj}. \quad (50)$$

Thus $\hat{H}$ is specified completely in terms of the known quantities in K. Rewriting (50) in matrix form yields

$$\hat{H} = K - \tilde{K}, \quad (51)$$

where

$$\tilde{K} = \begin{pmatrix} K_{11} & K_{22} & \cdots & K_{mm} \\ K_{11} & K_{22} & \cdots & K_{mm} \\ \vdots & \vdots & \vdots & \vdots \\ K_{11} & K_{22} & \cdots & K_{mm} \end{pmatrix}. \quad (52)$$

Thus, substituting for $\hat{H}$,

$$H - H_\Delta = K - \tilde{K}, \quad (53)$$

establishes a solution to the recurrence equation (25) for aperiodic irredicible Markov chains.

We now remove the assumption of aperiodicity. If P is periodic, P^n does not converge directly to A, but P^n is Cesaro-summable to A as in [5, p.101]. We proceed by summing (36) through (39) and dividing by n to give

$$\hat{H} = \frac{1}{n}\sum_{k=1}^{n} P^k\hat{H} + \frac{1}{n}\sum_{k=1}^{n}(I + P + P^2 + \cdots + P^{k-1})(H^* - H_\Delta). \quad (54)$$

We now use the relationship $A(H^* - H_\Delta) = 0$, as proved in (17) of Theorem 1, to write (54) in the equivalent form

$$\hat{H} = \frac{1}{n}\sum_{k=1}^{n} P^k\hat{H} + \frac{1}{n}\sum_{k=1}^{n}(I + P - A + P^2 - A + \cdots + P^{k-1} - A) \cdot (H^* - H_\Delta). \quad (55)$$

Again we apply $P^n - A = (P - A)^n$ from (41), to obtain

$$\hat{H} = \frac{1}{n}\sum_{k=1}^{n} P^k\hat{H} + \frac{1}{n}\sum_{k=1}^{n}\sum_{l=0}^{k-1}(P - A)^l(H^* - H_\Delta). \quad (56)$$

Since P^n is Cesaro-summable to A, we have

$$\frac{1}{n}\sum_{k=1}^{n} P^k \to A. \quad (57)$$

The inverse of $I - P + A$ exists [5, p.101] and is given by

$$\frac{1}{n}\sum_{k=1}^{n}\sum_{l=0}^{k-1}(P - A)^l \to (I - P + A)^{-1}. \quad (58)$$

Taking the limit as $n \to \infty$, we have

$$\hat{H} = A\hat{H} + (I - (P - A))^{-1}(H^* - H_\Delta), \quad (59)$$

which is precisely the expression in (44). From here the proof continues as for aperiodic chains. Thus, the solution to the recurrence relation (25) is established for irreducible periodic and aperiodic Markov chains.

We now finish the proof by proving the uniqueness of the solution H given in (27). Let H and H' be two solutions to the recurrence equation (25). Then,

$$H = H^* + PH - PH_\Delta \tag{60}$$
$$H' = H^* + PH' - PH'_\Delta. \tag{61}$$

By subtracting (61) from (60), we obtain

$$H - H' = P(H - H') - P(H_\Delta - H'_\Delta). \tag{62}$$

But from Theorem 1, it follows that $H_\Delta = H'_\Delta$, yielding

$$\begin{aligned} H - H' &= P(H - H') - P(H_\Delta - H'_\Delta) && (63)\\ &= P(H - H'). && (64)\end{aligned}$$

By iteration, we obtain

$$\begin{aligned} H - H' &= P(H - H') && (65)\\ &= P^2(H - H') && (66)\\ &\ \vdots && (67)\\ &= P^n(H - H'). && (68)\end{aligned}$$

Summing (65) through (68) and dividing by n yields

$$H - H' = \frac{1}{n}\sum_{l=1}^{n} P^l(H - H'). \tag{69}$$

Using the fact that P^n is Cesaro-summable to A, we obtain

$$H - H' = A(H - H'). \tag{70}$$

Now since each row of A is the vector of stationary probabilities μ, the columns of $(H - H')$ must have elements which are equal, and since the diagonal is zero, all elements of the columns of $H - H'$ must be equal to zero and thus $(H - H')$ must be zero. Thus $H = H'$ and the solution to (25) is unique. □

V. Conclusion

We can now calculate the matrix of trajectory entropies for the example in Section III. The irreducible Markov chain is depicted in Fig. 1 has transition matrix

$$P = \begin{pmatrix} 0 & 0.9 & 0 & 0.1 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{71}$$

From Section III, the matrix H_Δ is given by

$$H_\Delta = \begin{pmatrix} 5.32 & 0 & 0 & 0 \\ 0 & 2.80 & 0 & 0 \\ 0 & 0 & 5.07 & 0 \\ 0 & 0 & 0 & 9.25 \end{pmatrix}. \tag{72}$$

The matrices $K = (I - P + A)^{-1}(H^* - H_\Delta)$ and $\tilde{K}$ are respectively given by

$$K = \begin{pmatrix} -3.3564 & -0.0001 & 1.4995 & 0.5917 \\ 1.6436 & -0.5990 & 1.4337 & 0.9665 \\ 0.1436 & 0.7004 & -2.5991 & 1.7791 \\ 0.1436 & 0.7004 & -2.5991 & -7.4713 \end{pmatrix}, \tag{73}$$

and

$$\tilde{K} = \begin{pmatrix} -3.3564 & -0.5990 & -2.5991 & -7.4713 \\ -3.3564 & -0.5990 & -2.5991 & -7.4713 \\ -3.3564 & -0.5990 & -2.5991 & -7.4713 \\ -3.3564 & -0.5990 & -2.5991 & -7.4713 \end{pmatrix}. \tag{74}$$

Consequently, from Theorem 2, the matrix of trajectory entropies is

$$H = \begin{pmatrix} 5.32 & 0.60 & 4.10 & 8.06 \\ 5.00 & 2.80 & 4.03 & 8.44 \\ 3.50 & 1.30 & 5.07 & 9.25 \\ 3.50 & 1.30 & 0 & 9.25 \end{pmatrix}. \tag{75}$$

For example, the entropy of the trajectory from state 4 to state 2 is 1.30 bits. Note that the entropy of trajectory T_{43} is 0 bits, reflecting the (conditional) determinism of that path. Notice also that any path originating in state 4 goes directly to state 3 and thus has the same probability as a path which originates in state 3. Therefore, the entropy of a trajectory starting in state 4 and going to state 1 is the same as that of a trajectory starting in state 3 and going to state 1. Thus, $H_{31} = H_{41}$, which is verified by inspection of the matrix H of trajectory entropies.

In summary, H expresses the descriptive complexity of the trajectories.

Acknowledgment

The authors wish to thank P. Barbone for many helpful discussions.

References

[1] S. Lloyd and H. Pagels, "Complexity as thermodynamic depth," *Ann. Phys.*, vol. 188, pp. 186–213, 1988.

[2] S. Lloyd, "Valuable information," in *Complexity, Entropy, and the Physics of Information, SFI Studies in the Sciences of Complexity* (Ed. W. H. Zurek). Reading, MA: Addison-Wesley, 1990, pp. 193–198.

[3] C.H. Bennett, "Logical depth and physical complexity," in *The Universal Turing Machine: A Half-Century Survey* (R. Herken, Ed. Oxford: Oxford Univ. Press, 1988, pp. 227–257.

[4] L. Ekroot and T.M. Cover, "The entropy of a randomly stopped sequence," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1641–1644, Nov. 1991.

[5] J.G. Kemeny and J.L. Snell, *Finite Markov Chains*. New York: D. Van Nostrand, 1960.