Hierarchical coding and normal sequences.

A collection $K$ of words over a finite alphabet is called an S-code (or a code without overlapping) if, for words in $K$, (1) no word is a subword of another, (2) a prefix of a word cannot be a suffix. To any S-code $K$ the authors associate a Borel subshift $\Omega_K$ of the sequences obtained from the words in $K$ by concatenation and insertion.

It is shown that if not all words in $K$ have length one then there exists a unique measure of maximal entropy $\mu$ on $\Omega_K$. Actually $\mu$ is the image of a Bernoulli measure via a special kind of homomorphism called hierarchical coding. Some ideas of such a coding procedure go back to L. D. Meshalkin [Dokl. Akad. Nauk SSSR 128 (1959), 41–44; MR0110782], A. Kh. Zaslavski˘ı [Teor. Veroyatnost. i Primenen. 9 (1964), 318–326; MR0164377], and R. J. Grigorchuk and Stépin [in Ergodic theory and related topics, (Vitte, 1981), 207–229, Akademie-Verlag, Berlin, 1982; see MR0730763].

Next the authors construct a class of $\mu$-normal sequences by concatenating words of given lengths. A related construction for unilateral sequences was presented by A. Bertrand-Mathis [Ergodic Theory Dynam. Systems 8 (1988), no. 1, 35–51; MR0939059].

An analytic criterion of normality is also given in terms of the second derivative of the generating function

$$\phi_K(z) = \sum_{n=0}^{\infty} |\{B \in K: |B| = n\}| z^n.$$

{For the collection containing this paper see MR1359087}