TWO NOTES ON SUBSHIFTS

JOSEPH S. MILLER

Fix $n > 1$. For $S \subseteq n^{<\omega}$, we say that $X \in n^\omega$ *avoids* S if no $\sigma \in S$ is a substring of X. The class $\mathcal{Q}_S \subseteq n^\omega$ of all sequences that avoid S is called a *subshift* (or a *shift space*). See Lind and Marcus [3] for an introduction to subshifts. We prove two unrelated results about subshifts. In Section 1, we prove that every Medvedev degree that contains a Π^0_1 class contains, in fact, a Π^0_1 subshift. This answers a question from [5], where Simpson proves that every Π^0_1 Medvedev degree contains a 2-dimensional subshift of *finite type*, i.e., one for which the set of forbidden 2-dimensional words is finite. Note that nonempty 1-dimensional subshifts of finite type contain periodic sequences, so they are all Medvedev equivalent.

In Section 2, we give a condition on a set $S \subseteq n^{<\omega}$ that is sufficient to guarantee that the corresponding subshift $\mathcal{Q}_S$ is nonempty. We also give examples that satisfy the condition. The fact that there is a computable sequence of lengths such that any sequence of words with those lengths gives a nonempty subshift was proved by Cenzer, Dashti and King [1, Theorem 3.2]. We show, for example, that when $n = 2$ this sequence can be taken to be $5, 6, 7, \ldots$. As another application of the sufficient condition, we show that for any $d < 1$, there is a $b \in \omega$ and an infinite sequence such that if $\tau \in 2^{<\omega}$ is any of its substrings, then τ has prefix-free Kolmogorov complexly greater than $d\,|\tau| - b$.

1. Every Π^0_1 Medvedev degree contains a Π^0_1 subshift

If S is a computably enumerable (c.e.) set, then $\mathcal{Q}_S$ is a Π^0_1 class. Conversely, it is not hard to see that if $\mathcal{Q}$ is a Π^0_1 subshift, then the set S of all strings that appear in no element of $\mathcal{Q}$ is c.e. In this case, clearly $\mathcal{Q} = \mathcal{Q}_S$. We will show that, from a computability theoretic perspective, Π^0_1 subshifts can exhibit all of the behavior possible from arbitrary Π^0_1 subclasses of n^ω. For what follows, we restrict ourselves to binary sequences because elements of n^ω can easily be coded in binary.

Let $\mathcal{P}, \mathcal{Q} \subseteq 2^\omega$. We say that $\mathcal{P}$ is *Medvedev* (or *strongly*) reducible to $\mathcal{Q}$ and write $\mathcal{P} \leq_s \mathcal{Q}$ if there is a Turing functional Ψ such that if $A \in \mathcal{Q}$, then $\Psi(A) \in \mathcal{P}$ [4]. In other words, $\mathcal{P} \leq_s \mathcal{Q}$ if there is a uniform algorithm to transform any element of $\mathcal{Q}$ to an element of $\mathcal{P}$. Medvedev reducibility gives us a way to divorce the combinatorial structure of a class from the effective content of its members. As expected, we write $\mathcal{P} \equiv_s \mathcal{Q}$ to mean that $\mathcal{P} \leq_s \mathcal{Q}$ and $\mathcal{Q} \leq_s \mathcal{P}$ and call the equivalence classes *Medvedev degrees*.

Proposition 1.1. *If $\mathcal{P}$ is a Π^0_1 class, then there is a Π^0_1 subshift $\mathcal{Q}$ such that $\mathcal{P} \equiv_s \mathcal{Q}$.*

Date: February 23, 2009.

2000 *Mathematics Subject Classification.* Primary 03D30; Secondary 37B10, 68Q30.

The author was supported by the National Science Foundation under grants DMS-0601021 and DMS-065267, the latter being part of a Focused Research Group in Algorithmic Randomness.

Proof. The key feature of the coding we use is that a sequence $Y \in \mathcal{P}$ is coded by another sequence X in such a way that every tail of X also codes Y. First, we define collections of strings $\{a_\sigma\}_{\sigma\in 2^{<\omega}}$ and $\{b_\sigma\}_{\sigma\in 2^{<\omega}}$. Let $a_\lambda = 0$ and $b_\lambda = 1$, where λ represents the empty string. For $\sigma \in 2^{<\omega}$, let

$$a_{\sigma 0} = b_\sigma a_\sigma a_\sigma, \qquad b_{\sigma 0} = b_\sigma a_\sigma a_\sigma a_\sigma,$$
$$a_{\sigma 1} = a_\sigma b_\sigma b_\sigma \text{ and } \qquad b_{\sigma 1} = a_\sigma b_\sigma b_\sigma b_\sigma.$$

Define $S_\sigma = \{a_\sigma a_\sigma a_\sigma a_\sigma, b_\sigma b_\sigma b_\sigma b_\sigma, a_\sigma a_\sigma b_\sigma b_\sigma, b_\sigma b_\sigma a_\sigma a_\sigma, a_\sigma b_\sigma a_\sigma b_\sigma, b_\sigma a_\sigma b_\sigma a_\sigma\}$ and let $S = \bigcup_{\sigma\in 2^{<\omega}} S_\sigma$. It is not hard to see that if $X \in 2^\omega$ avoids S_λ, then (except for at most two initial bits) X is either formed from a_0 and b_0 or from a_1 and b_1. In the same way, we can see by induction that if X avoids S, then for each $n \in \omega$ there is a unique $\sigma \in 2^n$ such that X is formed from a_σ and b_σ (again, disregarding an initial segment shorter than a_σ). On the other hand, for any $Y \in 2^\omega$, the infinite sequence $\Psi(Y) = \bigcup_{n>0} a_{Y\upharpoonright n}$ avoids S. To see that $\Psi(Y)$ is well defined we observe that a_σ is a prefix of both $a_{\sigma 0}$ and $a_{\sigma 1}$, as long as $\sigma \neq \lambda$. The latter is immediate. For the former, note that $\sigma \neq \lambda$ implies that a_σ is a prefix of b_σ and hence of $a_{\sigma 0}$.

Let $W \subseteq 2^{<\omega}$ be a c.e. set of strings such that $\mathcal{P}$ is exactly the set of sequences with no initial segment in W. Let $T = S \cup \{a_\sigma \colon \sigma \in W\}$. Then T is a c.e. set of strings, so the induced subshift $\mathcal{Q}_T$ is a Π^0_1 class. Moreover, Ψ is an effective reduction of elements of $\mathcal{P}$ to elements of $\mathcal{Q}_T$, hence $\mathcal{Q}_T \leq_s \mathcal{P}$. For the other direction, let $\Phi(X) = \bigcup\{\sigma \in 2^{<\omega} \colon a_\sigma \text{ is a substring of } X\}$ and assume that $\Phi(X)$ stops converging as soon as an incompatibility is found. Clearly Φ is total on any X that avoids S. If we additionally assume that X avoids T, then $\Phi(X) \in \mathcal{P}$. Thus $\mathcal{P} \leq_s \mathcal{Q}_T$. □

2. A sufficient condition for a subshift to be nonempty

Proposition 2.1. *Let $S \subseteq n^{<\omega}$. If there is a $c \in (1/n, 1)$ such that*

$$\sum_{\tau\in S} c^{|\tau|} \leq nc - 1,$$

then there is an $X \in n^\omega$ that avoids S.

Proof. Fix $c \in (1/n, 1)$. Let $p = \sum_{\tau\in S} c^{|\tau|}$ and assume that $p \leq nc - 1$. Since $p < 1$, we know that S does not contain the empty string. For each $\sigma \in n^{<\omega}$, let $w(\sigma) = \sum\{c^{|\rho|} \colon (\exists\tau)\ |\rho| < |\tau| \text{ and } \sigma\rho \text{ ends in } \tau\}$. Think of $w(\sigma)$ as a measure of the pending threats to an extension of σ ending in a element of S. Note that $w(\lambda) = 0$, where λ is the empty string. Our goal is to build a sequence $X \in n^\omega$, one digit at a time, so that the weight stays below 1. Such an X avoids S because as long as $w(\sigma) < 1$, we know that σ itself does not end in an element of S. Say that we currently have σ such that $w(\sigma) < 1$. It is not hard to see that $c\sum_{0\leq i<n} w(\sigma i) = w(\sigma) + p < nc$, hence $\sum_{0\leq i<n} w(\sigma i) < n$. Therefore, $w(\sigma i) < 1$, for some $0 \leq i < n$. □

Corollary 2.2. *Assume that $S \subseteq n^{<\omega}$ contains at most one string of each length and let $L = \{|\sigma| \colon \sigma \in S\}$. If*

(*a*) $n = 2$ *and* $L \subseteq \{5, 6, 7, \dots\}$, *or*
(*b*) $n = 2$ *and* $L \subseteq \{4, 6, 8, \dots\}$, *or*
(*c*) $n = 3$ *and* $L \subseteq \{2, 3, 4, \dots\}$, *or*
(*d*) $n = 4$ *and* $L \subseteq \{1, 2, 3, \dots\}$,

then there is an $X \in n^\omega$ *that avoids* S.

Proof. In (a) and (b) we can apply the proposition with $c = \frac{\sqrt{5}-1}{2}$, the inverse of the golden ratio. For (c) and (d) we can use $c = 1/2$. Also note that (d) follows easily from (c) because a string of length 1 simply eliminates a digit. □

We finish with an application of Proposition 2.1 to effective randomness. We refer the reader to Li and Vitányi [2] for an introduction to *prefix(-free Kolmogorv) complexity*, K, which measures the complexity of finite binary strings. Schnorr proved that $X \in 2^\omega$ is *Martin-Löf random* if and only if $(\forall n)\ K(X \upharpoonright n) \geq n - O(1)$. Almost every infinite sequence is Martin-Löf random, so we know that almost every sequence's initial segments have high complexity. What about the complexity of substrings? Here things look different; almost every infinite sequence has arbitrary long substrings of zeros. Even so, the next result shows that we can find sequences such that every substring has *fairly* high complexity. We are limited by the fact that if X avoids *any* string $\tau \in 2^{<\omega}$, then $\limsup K(X \upharpoonright n)/n < 1$. In other words, the following result would fail for $d = 1$.

Corollary 2.3. *Let* $d < 1$. *There is an* $X \in 2^\omega$ *such that if* $\tau \in 2^{<\omega}$ *is a substring of* X, *then* $K(\tau) > d\,|\tau| - O(1)$.

Proof. Fix $d \in (0,1)$ and let $b = -\log(1-d) + 1$ (where log denotes the base 2 logarithm). Let $S = \{\tau \in 2^{<\omega} \colon K(\tau) \leq d\,|\tau| - b\}$. To apply Proposition 2.1, we let $c = 2^{-d}$. Then

$$\sum_{\tau \in S} c^{|\tau|} = \sum_{\tau \in S} 2^{-d,|\tau|} \leq \sum_{\tau \in S} 2^{-K(\tau)-b} \leq 2^{-b} \sum_{\tau \in 2^{<\omega}} 2^{-K(\tau)} \leq 2^{-b},$$

where the last step is Kraft's inequality. It is easy to show that $2^{-b} = (1-d)/2 < 2^{1-d} - 1 = 2c - 1$, for $d \in (0,1)$. □

References

[1] Douglas Cenzer, S. Ali Dashti, and Jonathan L. F. King. Effective symbolic dynamics. In Ruth Dillhage, Tanja Grubba, Andrea Sorbi, Klaus Weihrauch, and Ning Zhong, editors, *Proceedings of the Fourth International Conference on Computability and Complexity in Analysis (CCA 2007)*, volume 202 of *Electronic Notes in Theoretical Computer Science*, pages 89–99. Elsevier, 2008. CCA 2007, Siena, Italy, June 16–18, 2007.

[2] M. Li and P. Vitányi. *An introduction to Kolmogorov complexity and its applications*. Texts and Monographs in Computer Science. Springer-Verlag, New York, 1993.

[3] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.

[4] Yu. T. Medvedev. Degrees of difficulty of the mass problem. *Doklady Akademii Nauk SSSR*, 104:501–504, 1955.

[5] Stephen G. Simpson. Medvedev degrees of 2-dimensional subshifts of finite type. To appear in Ergodic Theory and Dynamical Systems.

Joseph S. Miller, Department of Mathematics, University of Connecticut, Storrs, CT 06269-3009, USA

E-mail address: joseph.miller@math.uconn.edu