Complexity for Probabilists
Homework 4

1. Consider the binary random sequence $(Y_0, Y_1, \ldots)$ obtained as follows: First let $(X_0, X_1, \ldots)$ be a Bernoulli($\frac{1}{2}$) sequence (i.e., $(X_0, X_1, \ldots)$ is i.i.d. with $\mathbf{Pr}(X_i = 0) = \mathbf{Pr}(X_i = 1) = \frac{1}{2}$), and then let

$$Y_i = \begin{cases} X_0 & \text{for } i = 0,1,2, \\ X_1 & \text{for } i = 3,4,5, \\ X_2 & \text{for } i = 6,7,8, \\ \vdots & \vdots \end{cases}$$

For instance, if $(X_0, X_1, \ldots) = (10010\cdots)$, then

$$(Y_0, Y_1, \ldots) = (111\,000\,000\,111\,000\cdots).$$

Show that with probability 1, the Kolmogorov complexity of initial segments of $(Y_0, Y_1, \ldots)$ satisfies

$$\lim_{n\to\infty} \frac{K(Y_0, \ldots, Y_{n-1})}{n} = \frac{1}{3}.$$

2. The definition of Martin-Löf randomness of a binary sequence $\mathbf{a} = (a_0, a_1, \ldots)$ needs to be digested carefully, before one can obtain a good feeling for it. Here it is again:

> An event $A \subseteq \{0,1\}^{\mathbf{N}}$ ($\mathbf{N}$ is, as usual, the set of natural numbers) is said to be a **computable almost sure property** if there exists an algorithm M, taking $(a_0, \ldots, a_n)$ and $k \in \mathbf{N}$ as input, outputting 1 (accept) or 0 (reject), and satisfying
>
> (i) $\mathbf{a} \in A$ if and only if there exists a K such that
>
> $$M((a_0, \ldots, a_n), k) = 1 \text{ for all } k \geq K \text{ and all } n.$$
>
> (ii) For all k and all $\mathbf{a} \in \{0,1\}^{\mathbf{N}}$,
>
> $$M((a_0), k) \geq M((a_0, a_1), k) \geq M((a_0, a_1, a_2), k) \geq \cdots.$$
>
> (iii) If $X = (X_0, X_1, \ldots)$ is a Bernoulli($\frac{1}{2}$) sequence, then for all k,
>
> $$\mathbf{Pr}\Big(\lim_{n\to\infty} M((X_0, \ldots, X_n), k) = 0\Big) \leq 2^{-k}.$$
>
> The sequence $\mathbf{a} \in \{0,1\}^{\mathbf{N}}$ is said to be **Martin-Löf random** if $\mathbf{a} \in A$ for all computable almost sure properties A.

Hence, if we want to show that a property $A \subseteq \{0,1\}^{\mathbf{N}}$ holds for all Martin-Löf random sequences, then it is enough to show that A is a computable almost sure property. Now do the following[1]:

[1] My idea about subproblems (a), (b) and (c) is that they are progressively more difficult. (a) is basically just an excercise in understanding the definition of computable almost sure properties. (b) is a (hopefully) reasonably simple example of a computable almost sure property, whereas (c) is a bit more technical. Subproblem (c) looks like it is a prerequisite for solving problem (d); however, if you fail to solve (c), then I would like you to go on with (d), pretending that you have already solved (c).

(a) Show that if $X = (X_0, X_1, \ldots)$ is a Bernoulli($\frac{1}{2}$) sequence, then

$$\mathbf{Pr}(X \text{ is Martin-Löf random}) = 1\,.$$

(b) Show that for every Martin-Löf random sequence $\mathbf{a}$, the pattern 10010 occurs somewhere in $\mathbf{a}$ (i.e., there exists an i such that $X_i = 1$, $X_{i+1} = 0$, $X_{i+2} = 0$, $X_{i+3} = 1$, $X_{i+4} = 0$).

(c) Show that for every Martin-Löf random sequence $\mathbf{a}$, we have

$$\limsup_{n\to\infty} \frac{a_0 + a_1 + \ldots a_{n-1}}{n} \leq 0.51\,.$$

(d) State a natural generalization of (c), and use it to show that for every Martin-Löf random sequence $\mathbf{a}$, we have

$$\lim_{n\to\infty} \frac{a_0 + a_1 + \ldots a_{n-1}}{n} = \frac{1}{2}$$

(this is the law of large numbers for Martin-Löf random sequences).