Substitution Sequences

It is helpful to know something about matrices to work on this problem.

Consider the substitution rule

\[ 0 \rightarrow 01, \quad 1 \rightarrow 0. \]

If one starts with a “seed” (finite string of zeros and ones) \( S_0 \) and applies this rule, one gets, for example:

\[
S_0 = 0, \quad S_1 = 01, \quad S_3 = 010, \quad S_4 = 01001, \quad S_5 = 01001010 \text{ etc.}
\]

or \( S_0 = 11, \quad S_1 = 00, \quad S_2 = 0101, \quad S_3 = 010010 \text{ etc.} \)

For some substitution rules, with the right starting point one iterates to an infinite “fixed point”, where each string \( S_k \) contains the previous one as a prefix. This happens above for starting value \( S_0 = 0 \).

A second example:

\[ 0 \rightarrow 01, \quad 1 \rightarrow 10 \]

This is a “constant length” substitution.

Possible Questions to Investigate

1. Can one determine how fast the lengths of the strings \( S_k \) increase with \( k \)? What allowable growth rates are there, for different substitution rules?

2. Which patterns of finite strings occur? Which don’t occur?

3. Does a particular substring occur with a “limiting density” in an infinite fixed point word? (Can a notion of density be well-defined?). For example, what is the density of occurrences of 0 in the first example above? Of occurrences of 1? Of the block 11?

4. (Kolakoski sequence) This is a sequence of 1’s and 2’s such that the length of consecutive terms of the same kind (either 1 or 2) reproduces the sequence.

It starts: 122112212211…

This has a block of length 1, followed by a block of length 2 (of 2’s), followed by another block of length 2 (of 1’s), followed by a block of length 1 etc. These block lengths are 1 22 1… and are supposed to reproduce the original word.

(a) Show that this word is uniquely defined, if we started with 1. What if the first symbol were 2?

(b) (Unsolved problem) Is there a limiting density for the occurrence of 1 in the limit word?
(c) (Unsolved problem) Which patterns occur in the limit word. For example 222 never occurs.

(d) Can the Kolakoski sequence be described as fixed point of a substitution? What if we allow a generalized rules that replaces a block with another block, for example 11 \rightarrow 2221.