Chomsky Normal Form CFG for $\{0^{1^\alpha}1^{2^\alpha}\cdots 0^{(k-1)^\alpha}1^{k^\alpha}\}$
Exposition by William Gasarch

Let $\alpha > 0$ be a constant. Some of the Big-O constants depend on α. All CFG's are assumed to be in Chomsky Nomral Form.

We leave out floors and ceilings that are needed to make quantitites natural numbers.

Let $n = 1^\alpha + \cdots + k^\alpha \sim \frac{k^{\alpha+1}}{\alpha+1} \sim k^{1+\alpha}$. Hence $k \sim n^{1/(\alpha+1)}$.

We present a CFG in for

$$\{0^{1^\alpha}1^{2^\alpha}\cdots 0^{(k-1)^\alpha}1^{k^\alpha}\}$$

Our grammar is in several parts

1. For every $1 \le j \le k$ with j odd we have a CFG G_j^0 with start symbol S_j^0 that generates 0^{j^α} that is of size $O(\log j)$. The size of all of these grammars is $\sum_{j=1,\ j \text{ odd}}^{k} O(\log j) = O(k \log k)$.

2. For every $1 \le j \le k$ with j even we have a CFG G_j^1 with start symbol S_j^1 that generates 1^{j^α} that is of size $O(\log j)$. The size of all of these grammars is $\sum_{j=1,\ j \text{ even}}^{k} O(\log j) = O(k \log k)$.

3. We construct a CFG for $\{0^{1^\alpha}1^{2^\alpha}\cdots 0^{(k-1)^\alpha}1^{k^\alpha}\}$.

   We want a CFG that generates $S_1^0 S_2^1 \cdots S_{k-1}^0 S_k^1$. This can be done with a CFG of size $O(k)$. We also put into the grammars all of the grammars mentioned aove. Hence the size of the CFG is

   $$O(\log k + k\log k) = O(k \log k) = O(n^{1/(\alpha+1)} \log n).$$

   If $\alpha = \frac{1}{10}$ then we get $O(n^{10/11} \log n)$. As α goes down the size of the CFG goes up but is is of the form $O(n^{1-\epsilon} \log n)$.

   So the Kolm way, or your construction, DO yield bigger grammars then can be gotten by a natural language.

   QUESTION: Is there a SMALLER CFG for the language above?